

```

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*
73: em_estp12:*
74: em_estp13:*
75: em_estp14:*
76: em_estp15:*
77: em_estp16:*
78: em_estp17:*
79: em_estp18:*
80: em_estp19:*
81: em_estp110:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

```

117: gb\_est48:\*  
 118: gb\_est49:\*  
 119: gb\_est50:\*  
 120: gb\_est51:\*  
 121: gb\_est52:\*  
 122: gb\_est53:\*  
 123: gb\_est54:\*  
 124: gb\_est55:\*  
 125: gb\_est56:\*  
 126: gb\_est57:\*  
 127: gb\_est58:\*  
 128: gb\_est59:\*  
 129: gb\_est60:\*  
 130: gb\_est61:\*  
 131: gb\_est62:\*  
 132: gb\_est63:\*  
 133: gb\_est64:\*  
 134: gb\_est65:\*  
 135: gb\_est66:\*  
 136: gb\_est67:\*  
 137: gb\_est68:\*  
 138: gb\_est69:\*  
 139: gb\_est70:\*  
 140: gb\_est71:\*  
 141: gb\_est72:\*  
 142: gb\_est73:\*  
 143: gb\_est74:\*  
 144: gb\_est75:\*  
 145: gb\_est76:\*  
 146: gb\_est77:\*  
 147: gb\_est78:\*  
 148: gb\_est79:\*  
 149: gb\_est80:\*  
 150: gb\_est81:\*  
 151: gb\_est82:\*  
 152: gb\_est83:\*  
 153: gb\_est84:\*  
 154: gb\_est85:\*  
 155: gb\_est86:\*  
 156: gb\_est87:\*  
 157: gb\_est88:\*  
 158: gb\_est89:\*  
 159: gb\_est90:\*  
 160: gb\_est91:\*  
 161: gb\_est92:\*  
 162: gb\_est93:\*  
 163: gb\_est94:\*  
 164: gb\_est95:\*  
 165: gb\_est96:\*  
 166: gb\_est97:\*  
 167: gb\_est98:\*  
 168: gb\_est99:\*  
 169: gb\_est100:\*  
 170: gb\_est101:\*  
 171: gb\_est102:\*  
 172: gb\_est103:\*  
 173: gb\_est104:\*  
 174: gb\_est105:\*  
 175: gb\_est106:\*  
 176: gb\_est107:\*  
 177: gb\_est108:\*  
 178: gb\_est109:\*  
 179: gb\_est110:\*  
 180: gb\_est111:\*  
 181: gb\_est112:\*  
 182: gb\_est113:\*  
 183: gb\_est114:\*  
 184: gb\_est115:\*  
 185: gb\_est116:\*  
 186: gb\_est117:\*  
 187: gb\_est118:\*  
 188: gb\_est119:\*  
 189: gb\_est120:\*

190: gb\_est121:\*  
 191: gb\_est122:\*  
 192: gb\_est123:\*  
 193: gb\_est124:\*  
 194: gb\_est125:\*  
 195: gb\_est126:\*  
 196: gb\_est127:\*  
 197: gb\_est128:\*  
 198: gb\_est129:\*  
 199: gb\_est130:\*  
 200: gb\_est131:\*  
 201: gb\_est132:\*  
 202: gb\_est133:\*  
 203: gb\_est134:\*  
 204: gb\_est135:\*  
 205: gb\_est136:\*  
 206: gb\_est137:\*  
 207: gb\_est138:\*  
 208: gb\_est139:\*  
 209: gb\_est140:\*  
 210: gb\_est141:\*  
 211: gb\_est142:\*  
 212: gb\_est143:\*  
 213: gb\_est144:\*  
 214: gb\_est145:\*  
 215: gb\_est146:\*  
 216: gb\_est147:\*  
 217: gb\_est148:\*  
 218: gb\_est149:\*  
 219: gb\_est150:\*  
 220: gb\_est151:\*  
 221: gb\_est152:\*  
 222: gb\_est153:\*  
 223: gb\_est154:\*  
 224: gb\_est155:\*  
 225: gb\_est156:\*  
 226: gb\_est157:\*  
 227: gb\_est158:\*  
 228: gb\_est159:\*  
 229: gb\_est160:\*  
 230: gb\_est161:\*  
 231: gb\_est162:\*  
 232: gb\_est163:\*  
 233: gb\_est164:\*  
 234: gb\_est165:\*  
 235: gb\_est166:\*  
 236: gb\_est167:\*  
 237: gb\_est168:\*  
 238: gb\_est169:\*  
 239: gb\_est170:\*  
 240: gb\_est171:\*  
 241: gb\_est172:\*  
 242: gb\_est173:\*  
 243: gb\_est174:\*  
 244: gb\_est175:\*  
 245: gb\_est176:\*  
 246: gb\_est177:\*  
 247: gb\_est178:\*  
 248: gb\_est179:\*  
 249: gb\_est180:\*  
 250: gb\_est181:\*  
 251: gb\_est182:\*  
 252: gb\_est183:\*  
 253: gb\_est184:\*  
 254: gb\_est185:\*  
 255: gb\_est186:\*  
 256: gb\_est187:\*  
 257: gb\_est188:\*  
 258: gb\_est189:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query %				DB	ID	Description
	Score	Match	Length				
C	1	122.8	3.6	1337	81	BF630719	BF630719 HVSMPEB001
	2	119.8	3.5	1280	80	BF264950	BF264950 HV_CEA001
	3	119.4	3.5	1101	219	CNS0021J	AL061936 Drosophill
	4	117.2	3.4	1101	219	CNS00EVL	AL069706 Drosophill
	5	114.4	3.3	1101	219	CNS0039G	AL063921 Drosophill
	6	110.8	3.2	1219	80	BF264728	BF264728 HV_CEA001
	7	104.8	3.1	829	221	CNS03LUB	AL249932 Tetraodon
	8	101.6	3.0	1101	219	CNS00EVL	AL069706 Drosophill
	9	101.4	3.0	1092	220	CNS020K7	AL175696 Tetraodon
	10	101.2	3.0	1101	219	CNS00EO7	AL069440 Drosophill
C	11	100.6	2.9	1021	219	CNS0014D	AL104032 Drosophill
	12	100.2	2.9	1101	219	CNS00LT2	AL078714 Drosophill
	13	99.8	2.9	997	219	CNS005TE	AL060767 Drosophill
C	14	99.8	2.9	1067	80	BF262585	BF262585 HV_CEA000
	15	99.6	2.9	1366	80	BF262464	BF262464 HV_CEA000
C	16	99.2	2.9	876	219	CNS009G1	AL053529 Drosophill
	17	98.4	2.9	1032	220	CNS020IP	AL206746 Tetraodon
	18	97.8	2.9	1135	220	CNS033GQ	AL226115 Tetraodon
	19	97	2.8	1092	220	CNS020K7	AL175696 Tetraodon
C	20	96.8	2.8	994	221	CNS04NOJ	AL298972 Tetraodon
	21	96.4	2.8	1101	219	CNS0021J	AL061936 Drosophill
C	22	95	2.8	974	219	CNS00ITT	AL075432 Drosophill
C	23	94.8	2.8	1101	219	CNS0176M	AL108152 Drosophill
C	24	94.4	2.8	1201	219	CNS013KE	AL106396 Drosophill
C	25	94.2	2.8	935	256	BF10881	BF10881 F24H6-Sp6.1
C	26	94.2	2.8	1106	80	BF264948	BF264948 HV_CEA001
C	27	94	2.7	661	220	CNS020DK	AL209800 Tetraodon
	28	94	2.7	945	221	CNS04D0J	AL285149 Tetraodon
C	29	94	2.7	1225	219	CNS0161D	AL106171 Drosophill
C	30	93.8	2.7	1101	219	CNS00EO7	AL069440 Drosophill
C	31	93.6	2.7	1101	219	CNS00EOY	AL069603 Drosophill
C	32	93.4	2.7	576	220	CNS035N7	AL228940 Tetraodon
C	33	93.4	2.7	848	245	AZ549808	AZ549808 ENTZF62TF
C	34	93.4	2.7	1190	220	CNS020N7	AL206908 Tetraodon
C	35	93.2	2.7	1007	80	BF2624865	BF2624865 HV_CEA001
C	36	93.2	2.7	1073	146	BF274651	BF274651 GA_EB002
C	37	93.2	2.7	1248	256	B11336	B11336 F19M10-Sp6
C	38	92.8	2.7	868	219	CNS010YV	AL099601 Drosophill
C	39	92.6	2.7	625	220	CNS036A2	AL229763 Tetraodon
C	40	92.6	2.7	524	220	CNS01U9O	AL167541 Tetraodon
C	41	92.2	2.7	1101	219	CNS003BD	AL064091 Drosophill
C	42	92.2	2.7	1101	219	CNS003DQ	AL064580 Drosophill
C	43	92	2.7	865	146	BF264686	BF264686 HV_CEA001
C	44	92	2.7	1101	219	CNS006TE	AL065901 Drosophill
C	45	91.8	2.7	1029	220	CNS01ZGM	AL174271 Tetraodon

## ALIGNMENTS

RESULT 1  
BF630719/C  
ID BF630719 standard: RNA: EST: 1337 BP.

XX	
AC	* BF630719;
XX	
SV	BF630719.1
XX	
DT	21-DEC-2000 (Rel. 66, Created)
DT	21-DEC-2000 (Rel. 66, Last updated, Version 1)
DE	HVSMEB0013H16f Hordeum vulgare seedling shoot EST library HVcDNA0002
DE	(Dehydration stress) Hordeum vulgare cDNA clone HVSMEB0013H16f, mRNA
DE	sequence.
XX	
KW	EST.
XX	

```
Db 780 AAATATATATAATA-----TAAATTAATAATACAAAAATTAATAATAATAATAATAATAA 727
Qy 1037 ttataaacaagaagtaataactaaacatacatgtgatcaaatcaatgacagatttga 1096
Db 726 AAAAAATTAATAATAATACATAATAATAATAATAATAATAATAATAATAATAATAATAA 667
Qy 1097 tcttaaatagaaatgattgagcaacctcaaaaatgtctctcttaggacacaaatctt 1156
Db 666 TAATAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 607
Qy 1157 tcttttagcttattaaagccggaggttcaactctctccctgtgagacttttggtttc 1216
Db 606 AAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 547
Qy 1217 aaatcttttttccaaaaataatagttagtttaattgagggcataatatttggtttaatt 1276
Db 546 AATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 487
Qy 1277 aagtcctagatttttttaggaccattctctaatcagcacaataatcctaaattgtaacaca 1336
Db 486 AATAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 427
Qy 1337 tttaaaacttaaaagtattgcaattcacaatcctttaaataatatatatatatata 1396
Db 426 TAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 367
Qy 1397 t---atatatatatatatgaagtatatagaacgataaactccttactcacaacat 1453
Db 366 TAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 307
Qy 1454 tagcccaaaaaaacatccatgcatttaactaggaatttttaacaaactcaaataggt 1513
Db 306 ATATAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 247
Qy 1514 tggtagttaaaaaaacaataatagtagtgatgtacatgcctttaaataatatatactca 1573
Db 246 TATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 187
Qy 1574 taccgaagtttttaattttgcaaatataatcacatttctctatcatcaattaaatacatt 1633
Db 186 AACCTTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 127
Qy 1634 aataatgcataattctgtaatatctatctatttcaattccatagacacaaacaaataa 1693
Db 126 ATCAGTAGCAGTACTCANTATCTAATAACAGAAAAATAGAAAAATAATAATAATAA 67
Qy 1694 acatatcaaatagtttttaacttaacaaaaa 1723
Db 66 ATATAAAAAATAATAACACCCACCATATAA 37
```

## RESULT 2

```
BF264950
ID BF264950 standard; RNA; EST: 1280 BP.
AC
XX BF264950;
SV BF264950.1
XX
DT 20-NOV-2000 (Rel. 65, Created)
DE 20-NOV-2000 (Rel. 65, Last updated, Version 1)
XX
DE HV_CEA0010N09f Hordeum vulgare seedling green leaf EST library HVCDNA0004
DE (Erysiphe infected & control) Hordeum vulgare cDNA clone HV_CEA0010N09f,
DE mRNA sequence.
XX
KW EST.
XX
OS Hordeum vulgare (barley)
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Triticeae; Hordeum.
XX
```

```

[1]
RN 1-1280
RP Wing R., Close T.J., Kleinhofs A., Wise R., Begum D., Frisch D., Yu Y.,
RA Anderson H., Dale J., Henry D., Kernodle S., Palmer M., Rambo T.,
RA Sasaki C., Schwartzbeck J., Simmons J., Choi D.W., Main D., Wood T.;
RT "Development of a genetically and physically anchored EST resource for
RT barley genomics";
RL Unpublished.
XX
CC Contact: Wing RA
CC Clemson University Genomics Institute
CC Clemson University
CC 100 Jordan Hall, Clemson, SC 29634, USA
CC Tel: 864 656 7288
CC Fax: 864 656 4293
CC Email: rwing@clemson.edu
CC Seq primer: AATTAAACCTCCTAAAGGG
CC High quality sequence start: 39
CC High quality sequence stop: 1197.
XX
Key Location/Qualifiers
source 1..1280
/db_xref="taxon:4513"
/db_xref="ESTLIB:6818"
/note="vector: lambdazap; Site_1: EcoRI; Site_2: XhoI"
/organism="Hordeum vulgare"
/cultivar="C116155 (M1613)"
/clone="HV_CEA0010N09f"
/clone_lib="Hordeum vulgare seedling green leaf EST library
HVCDNA0004 (Erysiphe infected & control)"
/tissue_type="seedling green leaf"
/lab_host="TJC121"
XX
SQ Sequence 1280 BP; 731 A; 15 C; 18 G; 353 T; 163 other;
```

```

Query Match 3.5%; Score 119.8; DB 80; Length 1280;
Best Local Similarity 43.9%; Pred. No. 2.6e-10;
Matches 471; Conservative 0; Mismatches 594; Indels 7; Gaps 4;
Qy 519 agaagcagaccagcaaatcttattgtgtatatacatatctataaagaagaagaattcatatt 578
Db 212 AANNANANANNNNNNANANANANANANNNNNNNNANANANANANANANANANANANANANN 271
Qy 579 actatagaattaaagctaaagcaggttgaaacgctgaaagcagaattctctaaataata 638
Db 272 NAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAATAAAAAATAAAT 331
Qy 639 gtaactgtcacaaacttattgtgtatatacatatctataaagaagaagaactcaaatata 698
Db 332 AATTAATAATTTAAAAATAAAAAATATAAAAAAATATATATATATAATATAANANANANAA 391
Qy 699 tgataaatcatcttaacaaaattctctatgaaattataataaaaaaagtcacatttgacac 758
Db 392 AAAAAATTTAATTAATTTANNNTTAAAAATTAANNAATTTAANAAAAAATAAAAAANAAAAA 451
Qy 759 ttaaaagggtgacataaacctctctccaaaaaaatacaaacatttataattctcaaa 818
Db 452 TTAANAAAAAANAAAAAANAAAAAATAAAAAANAAAAAANAAAAAANAAAAAANAAAAA 511
Qy 819 actatggtgtaatttctgtaaatcaaaaagaagaaggatttcttatcatcatttc 878
Db 512 AAAAAAANAAAAAANAAAAAANAAAAAATTAATTAATAAAAAAANAAAAAATAA 571
Qy 879 attattgtcaaaacttcaaatcttcatgtaatttgaaggaggaataataataat 938
Db 572 ATAAAAAATTTAAAAATAAAAAAANAAAAAATAAAAAATAAAAAATTAANNAATTTATT 631
Qy 939 gttgttttcttcttattgtttacatttcttggaatttatattcatcaacaaatgtat 998
Db 632 TTTTATTATTATATATATATATATATATATATATATATATATATATATATATATATATAT 687
```



QY 999 titaatcatgatgagattaccatcaaaagctgcaacttatataaaacaaagttaataac 1058  
 Db 688 TATAATTTTAAATATTTTAAATATTTTAAATATTTTAAATATTTTAAATATTTA 747  
 QY 1059 taacaatcatgatgacacatcaatcaatgacagttttgattccttaaaatagaatgattgag 1118  
 Db 748 TTAATTTTAAATATTTTAAATATTTTAAATATTTTAAATATTTTAAATATTTT 806  
 QY 1119 caaacctcaaaatgtctctcttaggatccacaaatctcttcccttagcttattaaagccgg 1178  
 Db 807 TATATTTTAAATATTTTAAATATTTTAAATATTTTAAATATTTTAAATATTTT 866  
 QY 1179 gattcacaactctctccctcttagactttttgcttttccaaatcttttcttccaaat 1238  
 Db 867 AAATTTTAAATATTTTAAATATTTTAAATATTTTAAATATTTTAAATATTTT 926  
 QY 1239 caataattagtttaaggcacaataattgttttgaatttaagtcacatgatttttaggac 1298  
 Db 927 AAATTTTAAATATTTTAAATATTTTAAATATTTTAAATATTTTAAATATTTT 986  
 QY 1299 catctcctcaacgacaaatctcctaaatgttaacacatttcaaaccttaaaagtgatgca 1358  
 Db 987 TAAAT 1046  
 QY 1359 ttcaaatcctcaaatatataatataatataatataatataatataatataatgaa 1418  
 Db 1047 TTAAAAATTTTATATATAAT-TAAATATAAAAAATTTAATTTAAATTTTAAATATA 1105  
 QY 1419 agttatagaacgacatcctctactcaacaaatagcccaaaacacatccataatgc 1478  
 Db 1106 TATAATATTTAATATATAAANAANAATATAAANAATATAAANAATATAAANAAT 1165  
 QY 1479 attaaactcagaaattttaacaaac-tcaaataggttggttagttaaataaaacaaatag 1537  
 Db 1166 ATATAATATAAANAANAATATAAANAANAANAANAANAANAANAANAANAANA 1225  
 QY 1538 tagatgcacatgcgtaccttttaaaatataatataatataatcgaagtttttaa 1589  
 Db 1226 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 1277

## RESULT 3

CNS0021J 1101 bp DNA GSS 03-JUN-1999  
 LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #  
 DEFINITION BAC05N11 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
 ACCESSION AL061936  
 VERSION AL061936.1 GI:4940214  
 KEYWORDS GSS.  
 SOURCE fruit fly.  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 REFERENCE 1 (bases 1 to 1101)  
 AUTHORS Direct Submission  
 .TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
 COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoss in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's PI and EST libraries. A more detailed description of the library

and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.bufile.edu/drosophila\\_bac.htm](http://bacpac.med.bufile.edu/drosophila_bac.htm).

## FEATURES

Location/Qualifiers  
 source  
 1. ..1101  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 /clone\_lib="RPCI-98"  
 /clone="BAC05N11"  
 /note="end : TET3"

BASE COUNT 631 a 7 c 28 g 289 t 146 others

## ORIGIN

Query Match 3.5%; Score 119.4; DB 219; Length 1101;  
 Best Local Similarity 39.8%; Pred. No. 3e-10;  
 Matches 437; Conservative 69; Mismatches 588; Indels 3; Gaps 2;

QY 400 ttittcttactactttatttcagactattgtctattttgctcacaactttttgatt 459  
 Db 2 TTTTNNNTTKTTTTTKKKTNTTNTTNTTGTGTTTNNNTTNTTNTTNTTNTTNTTNTT 61  
 QY 460 gtcaacttttcgatccctaaaggtttgacaaatttaccctgcctttttctccaaagaaatca 519  
 Db 62 TKTTTTTTTATTTTNNNAANAANKAAATNATAATTTTAAATTAATAAKKWDATA 121  
 QY 520 gaacgagccacagcaaatatttgattttcttataaaagaagaagaatcatcat 579  
 Db 122 ATAAAAAHHAAAAAAGGTGTTTCATGAAGTTTAAAAAAGTTTAAAAAAGTTTAAAAA 181  
 QY 580 cttatagaattaaagcgaagcagttgaaacgtgaaagcagaatttctcaaaagaaatag 639  
 Db 182 AANNTCNNNNNAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 241  
 QY 640 taaactgctcacaaacttattgtgtatataacatcatctataaaagaactcaaatat 699  
 Db 242 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 301  
 QY 700 gataatcatcttaacaaatctctatgaaattataataaaagaagtcacatttgacact 759  
 Db 302 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 361  
 QY 760 taaaaggttgacaaataaccgtctctccaaagaaatacaaacatttataattctaaag 819  
 Db 362 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 421  
 QY 820 ctatggtgtaattttgctgaatcacaagaagaagaggttctctatacataagtttca 879  
 Db 422 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 481  
 QY 880 ttattgatacaactttcaaatcttcaatgttaatttgaaagaaagaaataagataatag 939  
 Db 482 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 541  
 QY 940 ttgtttttgtttcttgatgacattttcaggaatataatattcatacaaaagaatgatt 999  
 Db 542 TTTTNTTTTNTTTTNTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 601  
 QY 1000 ttaaatgatgagagattaccatccaaaggtgcgaacttatataaaagaagtaataact 1059  
 Db 602 WTTTNTTAAAWAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 661  
 QY 1060 aaacaatcatgatacacaatcaatgacagttttgattcttaaaatagaatgattgagc 1119  
 Db 662 ANATAAANAATTTTAT 721  
 QY 1120 aaactcacaagaatgtctcttaggacacaaatcttcttcttagcttattaaagccgg 1179  
 Db 722 AAAATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 781  
 QY 1180 agttcaactctctccctctgtagactttttgttttcaatcttttcttcaaaatc 1239  
 Db 782 TTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 841





Matches 449; Conservative 0; Mismatches 522; Indels 18; Gaps 4;

[illegible]

RESULT	7
CNS03LUB	

[illegible]





	Query Match	2.98;	Score 100.6;	DB 219;	Length 1021;
	Best Local Similarity	40.88;	Pred. No. 3.7e-07;		
	Matches 414;	Conservative	78;	Mismatches 512;	Indels 10; Gaps 4;
Oy	725	atgaattataataaaaaagctcatttggacacttaaaagggtgacataaacgcgtctct	784		
	:       :	:         :	:	:       :	:
Db	1018	AASAAACAAAAAHHAAAHAHAAAAAAAAAWTTTTTTTTTTTAAACMAAAAWTTTTTTT	959		

RESULT	12
CNS00LT2	
LOCUS	CNS00LT2 1101 bp DNA GSS 14-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC: BRACH8P19 of RPECI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL078714



Dd    616 TWTTWTTTTTWWAATTTTTTWTAATTTTTTTTTTTAKTKKTTTTTTTTTTATT     675

QY    985 acacaaaaatgtattttaatatgatgagagtaccatccaaaaggctcgaaactataaa    1044  
      | | | | | | : | : | : | : | : | : | : | : | : | : | : | :  
Dd    676 AAATAAAAAAWTTTDWAAAWTTT T KKKKKAAAADKWDAKKWDGAKKAFTKTKDKKA    735

QY    1045 acaaatgaattaat 1055  
      | : | | : : : : :  
Dd    736 AWAADAADKKDRK 746

RESULT 13  
CNS005TE/c  
LOCUS  
DEFINITION  
  
ACCESSION AL060767.1 GI:4943573  
VERSION  
KEYWORDS fruit fly.  
SOURCE Drosophila melanogaster  
ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephyrroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 997)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT - Web : www.genoscope.cns.fr  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/The\_BDGP\_Drosophila\_melanogaster\_BAC\_library.html  
The BDGP library was prepared by Kazutoyo Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain yz; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.

FEATURES  
source Location/Qualifiers  
          1..997  
            /organism="Drosophila melanogaster"  
            /db\_xref="taxon:7227"  
            /clone\_lib="RPCI-98"  
            /clone="BACR12K22"  
            /note="end : TET3"

BASE COUNT     89 a     99 c     13 g     258 t     538 others

ORIGIN

Query Match                  2.9%; Score 99.8; DB 219; Length 997;  
Best Local Similarity       22.6%; Pred. No. 5e-07;  
Matches 180; Conservative 171; Mismatches 438; Indels     8; Gaps     3;

QY    7 aaagtcaccaaacgcgaatttagtaaaagaacctgattgaaattgaattgccgcaa 66  
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Dd    824 .ARRAGARARRRRARARAARARRRRRRRRRGARRRRRRGGRRGRGRRRRRR 765

QY    67 aatcgcgaagatatagagccgacgcggacaagtgaacacacacaaagcgcgtaaagatga 126  
      :: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Dd    764 RBAARARARAGARAAAARRRRRRARRAGAARRRRRRRRRRRRRRRGGARARRRRARRRA 705

QY    127 aacagtggggaaggaagagagaatcttaccgatcatcgcaggggaaagatgggaatcac 186  
      | : : : : : : | | | | : : | : : : : : : | : : : : : : : : : |  
Dd    704 ARAGAGAAARRRRRRRGGAARRRRGGRRGRGRRRRRGAGARRRRRRRMRTTRARRRRRRRGA 645





QY 1151 aatcttctcttagcttattaaagcggagtgacactctctctcccttagagacttttt 1210  
 |||||  
 Db 472 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 413  
 QY - 1211 gtttcaaatcttcttcttcaaaaaaacaataatagtagtgagcgaataatgttgg 1270  
 |||||  
 Db 412 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 353  
 QY 1271 ttaataagtcacatagatttttttaggacct 1301  
 |||||  
 Db 352 AGATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 322

RESULT 15  
 BF262464/c  
 ID BF262464 standard; RNA; EST; 1366 BP.  
 XX BF262464;  
 AC BF262464;  
 SV BF262464.1  
 XX  
 DT 20-NOV-2000 (Rel. 65, Created)  
 DT 20-NOV-2000 (Rel. 65, Last updated, Version 1)  
 XX  
 DE HV\_CEA0004Allif Hordeum vulgare seedling green leaf EST library HVcDNA0004  
 DE (Erysiphe infected & control) Hordeum vulgare cDNA clone HV\_CEA0004Allif,  
 DE mRNA sequence.  
 XX  
 KW EST.  
 XX  
 OS Hordeum vulgare (barley)  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Triticeae; Hordeum.  
 XX  
 RN [1]  
 RP 1-1366  
 RA Wing R., Close T.J., Kleinhofs A., Wise R., Begum D., Frisch D., Yu Y.,  
 RA Anderson H., Dale J., Henry D., Kernodle S., Palmer M., Rambo T.,  
 RA Saski C., Schwartzbeck J., Simmons J., Choi D.W., Main D., Wood T.;  
 RT "Development of a genetically and physically anchored EST resource for  
 RT barley genomics";  
 RL Unpublished.  
 XX  
 CC Contact: Wing RA  
 CC Clemson University Genomics Institute  
 CC 100 Jordan Hall, Clemson, SC 29634, USA  
 CC  
 CC Tel: 864 656 7288  
 CC Fax: 864 656 4293  
 CC Email: rwing@clemson.edu  
 CC Seq primer: AATTAACCTCACTAAAGGG  
 CC High quality sequence start: 428  
 CC High quality sequence stop: 598.  
 XX  
 FH Key  
 FH Location/Qualifiers  
 FT 1. .1366  
 FT /db\_xref="taxon:4513"  
 FT /db\_xref="ESTLIB:6818"  
 FT /note="Vector: lambdaZAP; Site\_1: EcoRI; Site\_2: XhoI"  
 FT /organism="Hordeum vulgare"  
 FT /cultivar="C116155 (M1a13)"  
 FT /clone="HV\_CEA0004Allif"  
 FT /clone\_lib="Hordeum vulgare seedling green leaf EST library  
 HVcDNA0004 (Erysiphe infected & control)"  
 FT /tissue\_type="seedling green leaf"  
 FT /lab\_host="TJC121"  
 XX  
 SQ Sequence 1366 BP; 260 A; 37 C; 0 G; 642 T; 427 other;

Query Match 2.9%; Score 99.6; DB 80; Length 1366;

Best Local Similarity 45.0%; Pred. No. 5.3e-07;  
 Matches 407; Conservative 0; Mismatches 491; Indels 6; Gaps 2;  
 QY 672 acatactataaagaacactcaaatatataatgataaatacatttttaacaaaattttctatgaata 731  
 |||||  
 Db 1323 AT 1264  
 QY 732 tataataaaaaagtcacttttgacacactaaaggtgacacataa-----cgcctctccc 786  
 |||||  
 Db 1263 AAAT 1204  
 QY 787 aaaaaaaatcaaacatttataattttcaaacatggtgtaattttgctggaatacaaa 846  
 |||||  
 Db 1203 AT 1144  
 QY 847 aaaaaagaagatttctatcataaagtttcatttattgtatcaaacctttcaaaatttccat 906  
 |||||  
 Db 1143 AATTTTAT 1084  
 QY 907 gtaattgaaaggaaaaaataagataataatgtttgtttttgtttcttctgttaccatttt 966  
 |||||  
 Db 1083 ATAANNITATGGTTAT 1024  
 QY 967 catggaatatatcatacaaaaaatgtattttaatatgatgagagattaccatccaa 1026  
 |||||  
 Db 1023 ATAAAAAT 964  
 QY 1027 aagtcgaacttatataaaa--caagttaataactaaacacacacatgcatgacacacaa 1085  
 |||||  
 Db 963 TATTTAGAT 904  
 QY 1086 gacagttttgacttaaatagaaatgattgagcaaacctcaaaaatgtcttctcttaggat 1145  
 |||||  
 Db 903 AATTAAT 844  
 QY 1146 cacaataatcttctcttagcttataaagcggagtgacactctctctccctgttagac 1205  
 |||||  
 Db 843 AAATAAGATTAAT 784  
 QY 1206 tttgttttcaaatcttcttcttcaaaaaatcaataatagtagtaattggtgagcacaataatt 1265  
 |||||  
 Db 783 TTTTAAATGATTAAT 724  
 QY 1266 tggtttaattagtcacatagattttttaggaccatctcttaacacgacaaatccctaa 1325  
 |||||  
 Db 723 NNATTTAT 664  
 QY 1326 attgtaacacatttaaacacttaaaagattgtgattcattccacaaatcccttaaatatata 1385  
 |||||  
 Db 663 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 604  
 QY 1386 atatatatatatatatatatatataatgaaagtttatagaaacgataactcccttac 1445  
 |||||  
 Db 603 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 544  
 QY 1446 tcaacaattagcccaaaaaaacatcccatgactttaactgaggaatttttaacaaactc 1505  
 |||||  
 Db 543 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 484  
 QY 1506 aataggttggttagttaaaaaaaacaaatagtagtagtgatcactgcgtaccttttaaaaa 1565  
 |||||  
 Db 483 AANAATAAT 1569  
 QY 1566 tata 1569  
 Db 423 ANNA 420

Search completed: July 12, 2001, 23:04:16  
 Job time: 4314 sec

---

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2001, 22:06:12 ; Search time 92.84 Seconds  
(without alignments)  
6832.523 Million cell updates/sec

Title: PCT-US01-13059-3  
Perfect score: 3424  
Sequence: 1 aagcttaagctaccacat.....aacggcatactgactcagt 3424

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 317530 seqs, 92630169 residues

Total number of hits satisfying chosen parameters: 635060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	115.6	3.4	19124	2	US-08-487-826B-13
C 2	103.8	3.0	7218	1	US-08-232-463-14
C 3	83.4	2.4	19124	2	US-08-487-826B-13
C 4	79.4	2.3	8920	2	US-08-446-855A-1
C 5	79.4	2.3	8920	4	US-09-150-741-1
C 6	71.8	2.1	6152	4	US-08-973-462-1
C 7	71	2.1	4673	1	US-07-638-431-1
C 8	70.6	2.1	4673	5	PCT-US92-00018-1
C 9	69.4	2.0	51952	4	US-07-867-106-2
C 10	68.4	2.0	837	4	US-08-947-823-1
C 11	68.4	2.0	837	4	US-08-998-416-288
C 12	67.6	2.0	615	4	US-08-998-416-186
C 13	66.2	1.9	3095	6	5231168-1
C 14	65.2	1.9	665	2	US-08-883-795A-36
C 15	64.8	1.9	665	2	US-08-883-795A-36
C 16	64.6	1.9	6152	4	US-08-973-462-1
C 17	64.2	1.9	6768	1	US-08-107-755A-1
C 18	64.2	1.9	8457	1	US-07-991-867B-1
C 19	64.2	1.9	8457	2	US-08-544-332-1
C 20	64.2	1.9	8920	2	US-08-446-855A-1
C 21	64.2	1.9	8920	4	US-09-150-741-1
C 22	63.8	1.9	636	4	US-08-998-416-1137
C 23	63.8	1.9	701	4	US-08-998-416-701
C 24	63.4	1.9	834	4	US-08-998-416-305
C 25	63	1.8	51952	4	US-08-947-823-1
C 26	62.6	1.8	6243	2	US-09-056-075-1
C 27	62.4	1.8	724	4	US-08-998-416-693

28	62.4	1.8	732	4	US-08-998-416-1036	Sequence 1036, Ap
29	62.4	1.8	837	4	US-08-998-416-288	Sequence 288, Ap
C 30	61.6	1.8	731	1	US-08-451-405A-2	Sequence 2, Appli
C 31	60.8	1.8	636	4	US-08-998-416-1137	Sequence 1137, Ap
C 32	60.6	1.8	1511	1	US-07-991-867B-8	Sequence 8, Appli
C 33	60.6	1.8	1511	1	US-08-107-755A-8	Sequence 8, Appli
C 34	60.6	1.8	1511	2	US-08-544-332-8	Sequence 8, Appli
C 35	59.8	1.7	1850	4	US-08-617-860B-32	Sequence 32, Appli
C 36	59.8	1.7	4098	2	US-08-605-106-4	Sequence 4, Appli
C 37	59.6	1.7	3095	6	5231168-1	Patent No. 5231168
C 38	59	1.7	658	4	US-08-998-416-595	Sequence 595, App
C 39	58.8	1.7	1368	3	US-08-874-563-5	Sequence 5, Appli
C 40	58.8	1.7	1368	4	US-08-577-483-14	Sequence 14, Appl
C 41	58.8	1.7	3701	4	US-08-845-258-10	Sequence 10, Appl
C 42	58.8	1.7	3701	4	US-08-990-571-10	Sequence 10, Appl
C 43	58.8	1.7	4253	4	US-08-577-483-7	Sequence 7, Appli
C 44	58.6	1.7	711	4	US-08-998-416-786	Sequence 786, App
C 45	58.6	1.7	782	4	US-08-998-416-224	Sequence 224, App

ALIGNMENTS

RESULT 1  
US-08-487-826B-13/C  
; Sequence 13, Application US/08487826B  
; Patent No. 5993827  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Kim L.  
; APPLICANT: Chitnis, Chetan  
; APPLICANT: Miller, Louis H.  
; APPLICANT: Peterson, David S.  
; APPLICANT: Su, Xin-zhaun  
; APPLICANT: Wellens, Thomas E.  
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Knobbe Martens Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: California  
; COUNTRY: US  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,826B  
; FILING DATE: 10-SEP-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: NIH121.001CPI  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19124 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-487-826B-13

Query Match 3.4%; Score 115.6; DB 2; Length 19124;



RESULT 2					
US-08-232-463-14/c					
; Sequence 14, Application US/08232463					
; Patent No. 5670367					
; GENERAL INFORMATION:					
; APPLICANT: DORNER, F.					
; APPLICANT: SCHEIFELINGER, F.					
; APPLICANT: FALKNER, F. G.					
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS					
; NUMBER OF SEQUENCES: 52					
; CORRESPONDENCE ADDRESS:					
; ADDRESSEE: Foley & Lardner					
; STREET: 1800 Diagonal Road, Suite 500					
; CITY: Alexandria					
; STATE: VA					
; COUNTRY: USA					
; ZIP: 22313-0299					
; COMPUTER READABLE FORM:					
; MEDIUM TYPE: Floppy disk					
; COMPUTER: IBM PC compatible					
; OPERATING SYSTEM: PC-DOS/MS-DOS					
; SOFTWARE: PatentIn Release #1.0, Version #1.25					
; CURRENT APPLICATION DATA: US/08/232.463					
; FILING DATE:					
; CLASSIFICATION: 435					
; PRIOR APPLICATION DATA:					
; APPLICATION NUMBER: US/07/935,313					
; FILING DATE:					
; APPLICATION NUMBER: EP 91 114 300.6					
; FILING DATE: 26-AUG-1991					
; ATTORNEY/AGENT INFORMATION:					
; NAME: BENT, Stephen A.					
; REGISTRATION NUMBER: 29,768					
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU					
; TELECOMMUNICATION INFORMATION:					
; TELEPHONE: (703)836-9300					
; TELEX: (703)683-4109					
; TELEX: 899149					
; INFORMATION FOR SEQ ID NO: 14:					
; SEQUENCE CHARACTERISTICS:					
; LENGTH: 7218 base pairs					
; TYPE: nucleic acid					
; STRANDEDNESS: single					
; TOPOLOGY: linear					
; IMMEDIATE SOURCE:					
; CLONE: pT2gpt-F1s					
US-08-232-463-14					
Query Match 3.0%; Score 103.8; DB 1; Length 7218;					
Best Local Similarity 2.5%; Pred. No. 1.5e-10;					
Matches 9; Conservative 257; Mismatches 99; Indels 0; Caps 0;					
QY	22	gaatttagtaaaagaccatgatatttgaaatcggaatgtcggcaaaaatcgagaagata 81			
DB	1445	GAATTTGTGTCACRRR 1386			
QY	82	gagccgcacgaggacaagtgaaccacaaagcgcgtaagaatgaacagctgggagaag 141			
DB	1305	RRR 1326			
QY	142	aagagagataattaccgatcatcttcagggaaaaaatgggaatcgagagaaaaatcgaaa 201			
DB	1325	RRR 1266			
QY	202	aaaagaaattaagagaagagagaagaagttagagagaagaactgagcagtgactgct 261			
DB	1265	RRR 1206			
QY	262	atagccacatcccacatggtgtgatgagagagagagagagagaggttaaagcacaaatt 321			

US-08-487-826B-13					
; Sequence 13, Application US/08487826B					
; Patent No. 5953827					
; GENERAL INFORMATION:					
; APPLICANT: Sim, Kim L.					
; APPLICANT: Chitnais, Chetan					
; APPLICANT: Miller, Louis H.					
; APPLICANT: Peterson, David S.					
; APPLICANT: Su, Xin-zhaun					
; APPLICANT: Wellemis, Thomas E.					
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX					
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS					
; NUMBER OF SEQUENCES: 45					
; CORRESPONDENCE ADDRESS:					
; ADDRESSEE: Knobbe Martens Olson & Bear					
; STREET: 620 Newport Center Drive 16th Floor					
; CITY: Newport Beach					
; STATE: California					
; COUNTRY: US					
; ZIP: 92660					
; COMPUTER READABLE FORM:					
; MEDIUM TYPE: Floppy disk					
; COMPUTER: IBM PC compatible					
; OPERATING SYSTEM: PC-DOS/MS-DOS					
; SOFTWARE: PatentIn Release #1.0, Version #1.25					
; CURRENT APPLICATION DATA:					
; APPLICATION NUMBER: US/08/487,826B					
; FILING DATE: 10-SEP-1993					
; CLASSIFICATION: 435					
; ATTORNEY/AGENT INFORMATION:					
; NAME: Israelson, Ned					
; REGISTRATION NUMBER: 29,655					
; REFERENCE/DOCKET NUMBER: NIH121.001CP1					
; TELECOMMUNICATION INFORMATION:					
; TELEPHONE: (619) 235-8550					
; TELEFAX: (619) 235-0176					
; INFORMATION FOR SEQ ID NO: 13:					
; SEQUENCE CHARACTERISTICS:					
; LENGTH: 19124 base pairs					
; TYPE: nucleic acid					
; STRANDEDNESS: single					
; TOPOLOGY: linear					
; MOLECULE TYPE: cDNA					
; HYPOTHETICAL: NO					
; ANTI-SENSE: NO					
US-08-487-826B-13					
Query Match 2.4%; Score 83.4; DB 2; Length 19124;					
Best Local Similarity 44.2%; Pred. No. 7.5e-07;					
Matches 345; Conservative 0; Mismatches 436; Indels 0; Gaps 0;					
QY	829	aatttgcgaatcaaaaaaagaagaaggattctatatcataagtttcatttgttat 888			
DB	15438	AAATTATTAAAAATGAAAAAAGAAAAATCAATAATAAAAAAAATTTATTAATAAAAA 15497			
QY	889	cnaacttccaatttcattgttaatttgaaaggaaaaataataatgattgttttg 948			
DB	15498	AAAAAANAANAANAAGAGAAAAATTTTTTAAAAAATAATAATAATAATAATAATA 15557			





EARLIER APPLICATION NUMBER: PL6380  
EARLIER FILING DATE: 1992-12-16  
EARLIER APPLICATION NUMBER: AU93/00617  
EARLIER FILING DATE: 1993-12-02  
EARLIER APPLICATION NUMBER: 08/446,855  
EARLIER FILING DATE: 1995-07-06  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 8920  
TYPE: DNA  
ORGANISM: Plasmodium falciparum  
US-09-150-741-1

Query Match 2.3%; Score 79.4; DB 4; Length 8920;  
Best Local Similarity 48.7%; Pred. No. 3.5e-06;  
Matches 248; Conservative 0; Mismatches 256; Indels 5; Gaps 1;  
QY 1232 aaaaaataaataattgtaattggaatggaatatttgggttttaattgaagtcacatagatttt 1291  
Db 165 ataaaaataaaagaataaaaaatttatatttgatattgtagtatataatgattat 224  
QY 1292 ttaggaccatcttaacgcagcaaatatcctaattgtaaacacatttaaaacttaaaag 1351  
Db 225 tcaataataacacatagataaaaaacttttttttttttttttttttttttttttttttt 284  
QY 1352 tatggaticacaaactcttaaaatatatatatatatatatatatatatatatatatat 1411  
Db 285 caatacatattgaatttttatatatatatatatatatatatatatatatatatatat 344  
QY 1412 atatagaagt-----tatagaacgcataactccttactcaacaattagcccaaaaaa 1466  
Db 345 agtttgggtgttcattgtttataaaattacttgaaataataaaaaacttataatatatt 404  
QY 1467 catccataatgcatttaaaactaggaattttaacaaactcaaataggttggttagttaa 1526  
Db 405 ccaattcaatgaatacaattataatttattgattgtgacacatttaattagttttaca 464  
QY 1527 aaacaaaatagtagtagtgcatacgtaccccttaaaataataataactcatatcagaa 1586  
Db 465 ctcttataataaaacacatcctatatatatatacacacataataataactcccaattgt 524  
QY 1587 aaatttgcgaataataacatttatctatcaattaaataacatttaataatgcataat 1646  
Db 525 ggctccataaatttttatatatattttattttatttttttttttttttttttttttt 584  
QY 1647 tctgtaatatctcttttaatttccatatagaaacccaaacaaataaaacatatcaaatag 1706  
Db 585 cttagtttataaaatagtaattctactactaatttaaaaaaaataaaaaaaataaaaaa 644  
QY 1707 ttttaacttaacaaaacggttagggaaa 1735  
Db 645 agaaaaaaataatttaccatatgaaaa 673

RESULT 6  
US-08-973-462-1  
Sequence 1, Application US/08973462B  
Patent No. 6191270  
GENERAL INFORMATION:  
APPLICANT: DRUILHE, PIERRE  
APPLICANT: DAUBERSTES, PIERRE  
TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES  
FILE REFERENCE: 0660-0125-0 PCT  
CURRENT APPLICATION NUMBER: US/08/973,462B  
CURRENT FILING DATE: 1998-02-06  
EARLIER APPLICATION NUMBER: PCT/FR96/00894  
EARLIER FILING DATE: 1996-06-12  
EARLIER APPLICATION NUMBER: FR 95/07007  
EARLIER FILING DATE: 1995-06-13  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1  
LENGTH: 6152  
TYPE: DNA  
ORGANISM: P. falciparum  
US-08-973-462-1  
Query Match 2.1%; Score 71.8; DB 4; Length 6152;  
Best Local Similarity 47.9%; Pred. No. 7.6e-05;  
Matches 280; Conservative 0; Mismatches 292; Indels 13; Gaps 2;  
QY 509 aagaaaaatcagaacagacacagcaaaatttattgtatttttttattataaaagaaag 568  
Db 536 aataaaaaataaggagagaccgtttttattcatttctgtattgtatttttaagaattcaaaa 5425  
QY 569 aattcatattacttatagataattaaaagctaaagcagttgaaacgtgaaagcagaatttct 628  
Db 5426 cattattacaacaataaagaagaaatttagtaagaaaaataataatgaatagaa 5485  
QY 629 aaaaaaataagtaaacctgtacaaaactttatttgatgatatataacatactataaagaaa 688  
Db 5486 gtaacaaacaaagctgagaaaaaggtatgtacaggttaacaaataaaacccgagaaaaa 5545  
QY 689 ctcaaatatatgataaatcatttttaacaaaattttctgaaattataataaaaaaagtca 748  
Db 5546 actaaagttgataaaaataat-----aaagtaccgaaaaaaagaaagcgaataatca 5599  
QY 749 cttttgacacttaaaaggttgacaataacgctctctccaaaaaaataacaaacatttat 808  
Db 5600 aaataaaaaatgcagaagagtgaaatgattggagcgacaataaaataatgaatgaataaa 5659  
QY 809 aatttcaaaactatggtgtaattttgtctgaaatcaaaaagaaagaaagatttctatat 868  
Db 5660 aatataaaaaatgatatatttatgtaaatatataataataataataataataacatacat 5719  
QY 869 cataagtttcattattgtatcaaaactttcaaatctcatgtaatttgaagggaaaaaatt 928  
Db 5720 atatatatatatatgtattctttttacaaaattttcaaaattttcaaaattttatatatt 5779  
QY 929 aggataataatgtttgttt 988  
Db 5780 aatattta-----tattttccatatataaattttatttttcaatttttttttaatt 5832  
QY 989 aaaaatgtatttttaatatgatgagagattaccatcccaaaaggctgaacttatataaaaa 1048  
Db 5833 ataaatgttttttacagagtttatgttttttttaataatataatagatttctgtgaagaact 5892  
QY 1049 agttaataactaaacatacatgtgatacaatcaatgacagttt 1093  
Db 5893 gtatatattcatcacgatatatgtaataattattttgtgtttt 5937  
RESULT 7  
US-07-638-431-1/C  
Sequence 1, Application US/07638431  
Patent No. 5198535  
GENERAL INFORMATION:  
APPLICANT: Hoffman, Stephen L.  
APPLICANT: Charoenvit, Yupin  
APPLICANT: Hedstrom, Richard  
APPLICANT: Khumsmith, Seisin  
APPLICANT: Rogers IV, William O.  
TITLE OF INVENTION: Protective malaria sporozoite surface protein  
FILE REFERENCE: immunogen and gene  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: A. David Spevack  
STREET: NMRDC Building 1 T-12 National Naval  
CITY: Bethesda  
STATE: MD  
COUNTRY: USA  
ZIP: 20814-5044

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/638,431
FILING DATE: 19910110
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Spevack, Avrom D.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 295-6759
TELEFAX: (301) 295-4033
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4673 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: N
ANTI-SENSE: N
ORIGINAL SOURCE:
ORGANISM: Plasmodium yoelii
STRAIN: 17X(NL)
DEVELOPMENTAL STAGE: erythrocytic stage
TISSUE TYPE: Blood
CELL TYPE: erythrocytic stage
IMMEDIATE SOURCE:
LIBRARY: Py-lambda gtl1-2-7 kb genomic expression
CLONE: Py10.1111
FEATURE:
NAME/KEY: CDS
LOCATION: 718..3195
OTHER INFORMATION:
US-07-638-431-1

Query Match          2.1%; Score 71; DB 1; Length 4673;
Best Local Similarity 44.4%; Pred. No. 0.0001;
Matches 530; Conservative 0; Mismatches 640; Indels 25; Gaps 5;

Qy 510 agaaaaatcagaacagaccagcaaatattatgtattttctattataaaagaagaaga 569
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4563 AAACGAACTAGAAGGACATTATAAATTTATTCACATTATATGATCTTCAAATTA 4504

Qy 570 attcatattactatagaattaaagctaaagctgaaacggtgaaagcagaatttcta 629
      || || || || || || || || || || || || || || || || || || ||
Db 4503 AAAGTAGAGGTAATGGCGTTATTAGAAATTTATTCGATTTTATATAAAGGNAAT 4444

Qy 630 aaaaaatagtaaaactgtacaaactttattgtgtatataacatatctataagaac 689
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4443 AAATAAATGATTGATACAAATAGAAATATTTATTCCTTCTTCTTCTTCTTCTT 4384

Qy 690 tcaatatatgataaacttttcaaaaaatttctatgaaattataataaaaaaagtcac 749
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4383 AAAAATGATGAATATATATACAACTCAAAATTAATAAATTTACAAAATAAGAA 4324

Qy 750 ttgtacacttaaaaggttgacaataaccgctctctccaaaaaaatacaaacattata 809
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4325 ---TTAAGCAATAAAGAAATTAATAAATTTATATGCTCTCAAAACAAAATGG 4269

Qy 810 atttctaaactatggtgaattttgctgaaatcaaaaaaagaagaagatttctatc 869
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4268 ACTACAAATATGAGAAATGAAAAATTAATGGAATGATGATGCTTCTTCTTCTT 4209

Qy 870 ataagtttcatttattgtatcaaacctttcaaatttctatgaaattgaaaggaaata 929
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4208 AAAATTAATTTAATGATTTAAAAAATAAATCGATCTTCTTCTTCTTCTTCTT 4149

Qy 930 agataaatgtgtttttgtttcttctgtacatttttcacggaatatatttcatacaa 989
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 8
PCT-US92-00018-1/C
Sequence 1, Application PC/YUS9200018
GENERAL INFORMATION:
APPLICANT: Hoffman, Stephen L.
APPLICANT: Charoenvit, Yupin
APPLICANT: Hedstrom, Richard
APPLICANT: Khumsmith, Srisin
APPLICANT: Rogers IV, William O.
TITLE OF INVENTION: Protective malaria sporozoite surface protein
TITLE OF INVENTION: Immunogen and gene encoding
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: A. David Spevack
STREET: NWRDC Building 1 T-12 National Naval
CITY: Bethesda
STATE: MD
COUNTRY: USA
ZIP: 20814-5044
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/00018
; FILING DATE: 19920103
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Spevack, Avram D.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 295-6759
; TELEFAX: (301) 295-4033
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4673 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium yoelii
; STRAIN: 17X(NL)
; DEVELOPMENTAL STAGE: erythrocytic stage
; TISSUE TYPE: Blood
; CELL TYPE: erythrocytic stage
; IMMEDIATE SOURCE:
; LIBRARY: Py-lambda gt11-2-7 kb genomic expression
; CLONE: Py10.1111
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 718..3195
; OTHER INFORMATION:
;
; PCT-US92-00018-1

```

```

Query Match      2.1%; Score 71; DB 5; Length 4673;
Best Local Similarity 44.4%; Pred. No. 0.0001;
Matches 530; Conservative 0; Mismatches 640; Indels 25; Gaps 5;

QY 510 agaaaaatcagacagacacagcaaatatgtatttcttataaaagaagaaga 569
Db 4563 AACGAACTAGAGGAGACATATAAATTTATTCACATTATATGACTTCAAAATTATGA 4504

QY 570 attcatattactatagaattaaagcgaagcgttgaaagcgtgaagcagaatttcta 629
Db 4503 AAAGGTAGAGGTAATGCGCTTATTAGAAATTTATTTCGATTTTATAAAGGAAATTA 4444

QY 630 aaaaaaatagtaaaactgctacaaacttatttattgtgtatataacatatctataaagaac 689
Db 4443 AAATAAATGATTTGATACAAATAAAGAAATATATTTCCTTCTTAATAATATGAACAGAC 4384

QY 690 tcaaatatataataatcttttaacaaatttctatgaattataataaaaaaagtcac 749
Db 4383 AAAATGGGTAATATATATACAAATCTCAAAATTAATAAATTTACAAAATAAAGAA-- 4324

QY 750 tttagacacttaaaaggttgacaaataacgctctctccaaaaaaatcaaaacatttata 809
Db 4325 ---TTAAGCAATAAAGAAATTAATAAATTTATGCTTCCAAACAAAATAAAGGACATGCA 4269

QY 810 attctaaactatggtgttaatttctgtgaaatcaaaagaagaagagatttcttatac 869
Db 4268 ACTACAAATAAATGAAGAAATGAATAAATTTGGAATAATGATATTCTCTCTTAATAAAT 4209

QY 870 ataagttcattattgtatcaaaacttcaaatctcaatttcaatttgaagaagaagaattta 929
Db 4208 AAAATTAATTTTAAATGATTAAATAAATAAATAATTCGATTTCTTAATTTTGTAAAAATCAAA 4149

QY 930 agataataagttgttttcttctgttctgttcttctgttctgttctgttctgttctgttctgt 989
Db 4148 GAGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4089

```

```

QY 990 aaaaatgtattttaaataatgatgatgagagattaccatcccaaaagggtogaacttat-----ata 1043
Db 4088 GACAAATTTGTCAAAATATGCTAAATAATATATCACTCTTAATTTTACACAAATAATATAATA 4029

QY 1044 aaacaagtttaataactaaacaatacatagatgacacacacacacacacacacacacacacacac 1103
Db 4028 AATGAATCCAAACATATAACACAGATATGAANAATCAAAATATGAATTTTTCGAAAAGT 3969

QY 1104 atagaaatgatgagcaaacctcaaaaatgtctctcttaggatcacacaaaactcttcccttla 1163
Db 3968 TTATATGAAAATAATTAATAATTTTCAGATATGATAAATGAAGATGATAACTTCAAAAAAT 3909

QY 1164 gattattaaagccggaggttcaactctctctccctctgtgagactcttctgttcttcaaatctt 1223
Db 3908 CATGAGCCAACTTTAGATATTTCTTTTAAACATCCCATCATTTTCTTTTATACACTT 3849

QY 1224 ttctttcaaaaaatcaataatagtggaagggcacaataattgtttttaaataagtcac 1283
Db 3848 TTTAGTTCATAAACTTAAGACCATTTATTTTATGTTTGTAAATTTTAAATTAATACACA 3789

QY 1284 tagatttttttaggaccatctctaatacagacacaaatccctaaatctgtaacacacattttaa 1343
Db 3788 TTTTGTGTTAT-----TTCCTTTTATCGATAATAATTGGTGGATTTGCTCTATTTTGTAG 3735

QY 1344 cttaaagattatgacattcccaactccttaaaatatatatatatatatatatatatatat 1403
Db 3734 GAATTCATTTTATATGATATATC-----ACTTTTTTGTGTTTATCATTAATTTTGTGA 3680

QY 1404 atatatatatgaagattatagaaacgataaactccttactcaacaacatagcccacaa 1463
Db 3679 AATAGTAATACCGTGTAAATATACAACTTAAANAATGTTATTAACCTTTTAACTTTTAACTTTT 3620

QY 1464 aaacatcccaataatgcatcttaaaactaggaatttttaacaaactcaaaataggtgttagtta 1523
Db 3619 TTTTCTTTTCTTTT---TTTATATTAAGAATAATAATTCTAACCATTAATAATTTGAGTATA 3563

QY 1524 aaaaaaataatgtagatgacacacgctacacacacacacacacacacacacacacacacacacac 1583
Db 3562 ATAAATATATATATTAATACGACGACAAATTAGTTAAAAAATAAATAAATAAATAAATAAATAA 3503

QY 1584 tttaaatgttggaataataacacattctctatcaataataataataataataataataataatgcat 1643
Db 3502 TAAAAAATAATGAAATAAATAATGATATATAATTCGAATGAATAAATAAATAAATAAATAAATAA 3443

QY 1644 aattctgtatctatcttcttatttcccatatagacacacacacacacacacacacacacacacac 1698
Db 3442 ATTTATCAAACTGTTAATTTAAATAAATTAATCAATCAACAAAAAATAAATAAATAAATAAATAA 3388

RESULT 9
US-07-867-106-2/c
; Sequence 2, Application US/07867106
; Patent No. 5389526
; GENERAL INFORMATION:
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Keith L
; TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostellium
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
; STREET: One Liberty place 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/07/867,106  
 FILING DATE: 19920625  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PU PJ 7187  
 APPLICATION NUMBER: PCT/AU90/00530  
 FILING DATE: 02-NOV-1989  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Feeney, Joanne Longo  
 REGISTRATION NUMBER: 35,134  
 REFERENCE/DOCKET NUMBER: RICE-0002  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 215-568-3100  
 TELEFAX: 215-568-3439  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 5852 base pairs  
 TYPE: NUCLEIC ACID  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 ANTI-SENSE: NO  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 2378..5038  
 NAME/KEY: CDS  
 LOCATION: 2378..5038  
 US-07-867-106-2

Query Match 2.1%; Score 70.6; DB 1; Length 5852;  
 Best Local Similarity 42.8%; Pred. No. 0.00012;  
 Matches 523; Conservative 0; Mismatches 689; Indels 9; Gaps 3;

Qy 537 tttatgtattttcttataaaagaagaagaattctattacttatagattataaaagc 596  
 Db 2443 TTTGAATTTTCCAAAAGTATTACAAAACCTTAAACAACTATCCCAAGAAATAAGTTC 2384  
 Qy 597 taagcagtgaaacgtgaagcagaattcttaaaaaaaatagtaaacgtctcaaaactt 656  
 Db 2383 GTCCATTTTAAATCTTTGAATAAATTAACCAAACTATTAAATTAATAAATAATAA 2324  
 Qy 657 attatgtgtataaocattctataagaacactcaaatatataatgataaatttttaaca 716  
 Db 2323 AATGAATTTACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2264  
 Qy 717 aaattctatgaattataataaaaaaagtcacitttgacacttaaaaggttgacaataa 776  
 Db 2263 AATTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2204  
 Qy 777 ccgtctctccaaaaaaatcaaacatttataatttctaaacactatggtgtatatttgc 836  
 Db 2203 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2144  
 Qy 837 tgaatcaaaaaagaagaagattctctatcaatgaatttcaattatgtatcaaaacttt 896  
 Db 2143 AATATTTAATAAAGGTATTAATAGATAGAGAATAAATAAATAAATAAATAAATAAATAA 2087  
 Qy 897 caaatctcatgtattgaaagaaaaaataagataatgttatttatttatttatttattt 956  
 Db 2086 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2027  
 Qy 957 gttacattttcatggaatataattctatacaaaaaatgtatttataatgatgatgat 1016  
 Db 2026 TTTTATTTTTCAGAGTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1967  
 Qy 1017 taccatcaaaaggtgcgaacttatataaaacaggttaataaactaaacatacatgtatc 1076  
 Db 1966 GGTAAACATACATTAGTTTTTATAGTTTTTGCATATTAAAAATAAATAAATAAATAAATAA 1907  
 Qy 1077 acaatcaatgacagttttgtatttataaataagaaatgttgacgaacacacataatgtct 1136  
 Db 1906 ATTGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1847

Qy 1137 tcttaggatcacaaatc---tttctttagcttattaaagccggagttcaactctctc 1193  
 Db 1846 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1787  
 Qy 1194 tcccttgtagacttttggtttccaaatcttttcttccaaataatcaataatgttgaat 1253  
 Db 1786 AAGTATATATCATGATAGGCAATTTATTTATATATATATATATATATATATATATATAT 1727  
 Qy 1254 gggcataatatttggttttaataagccatagatttttttaggaccatctctaatcacga 1313  
 Db 1726 AATGTCATCAATAGTATTTTAAACATTTTATTTTAAAAAATAAATAAATAAATAAATAA 1667  
 Qy 1314 caaatatccttaaatgttaaacacatttaaaagttaagattgcttaccatccttaaa 1373  
 Db 1666 CAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1607  
 Qy 1374 at 1433  
 Db 1606 CTAGGTTAGTTTTTATATAATTTTACATATTTGTTAAATACTTTTAAATTTGAATCATAT 1547  
 Qy 1434 ataactcttactcaacaatttagcccaaaaaaacatccataatgcatttaaaactaggaat 1493  
 Db 1546 GATATTACATCGTCCGTTGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1487  
 Qy 1494 tttaacaaactcaaataggttggttagtttaaaaaaa---aacaaatagtagatgtacatac 1550  
 Db 1486 TTTTAAAAAATGATATAAATAATTTTAACTAACTATTTTAAATAAATAAATAAATAAATAA 1427  
 Qy 1551 gtacctttaaaataataactatcatcatcgaaagttttaaaattttgcgaaatttaaacatt 1610  
 Db 1426 TTATCTTAATCAATTTTGGTTTATACATATTTATGTTCTGACTGAAGTATAGATCTT 1367  
 Qy 1611 tatctataatataaataacatttaataatgcataattctgtaataatctctatcttatttc 1670  
 Db 1366 ATTACTAAAGTTTCAAAAGTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1307  
 Qy 1671 catatgaac 1730  
 Db 1306 TTTTTCATATCTCTCATGACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1247  
 Qy 1731 gaaaagttgaccttaactagct 1751  
 Db 1246 AAATTTTATTTCTCAATAGGT 1226

RESULT 10  
 US-08-947-823-1/c  
 : Sequence 1, Application US/08947823  
 : Patent No. 6114605  
 : GENERAL INFORMATION:  
 : APPLICANT: Williamson, Valerie M.  
 : APPLICANT: Kaloshian, Isgouhi  
 : APPLICANT: Yaghoobi, Jafar  
 : APPLICANT: Bodeau, John  
 : APPLICANT: Milligan, Stephen  
 : TITLE OF INVENTION: Procedures and Materials for Conferring  
 : PEST RESISTANCE IN PLANTS  
 : NUMBER OF SEQUENCES: 5  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Townsend and Townsend and Crew LLP  
 : STREET: Two Embarcadero Center, Eighth Floor  
 : CITY: San Francisco  
 : STATE: California  
 : COUNTRY: USA  
 : ZIP: 94111-3834  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/947,823

```

; FILING DATE: 09-OCT-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/18802
; FILING DATE: 09-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/028,191
; FILING DATE: 10-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REFERENCE/DOCKET NUMBER: 023070-070210US
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 51952 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-947-823-1

Query Match 2.0%; Score 69.4; DB 4; Length 51952;
Best Local Similarity 43.4%; Pred. No. 0.00027;
Matches 523; Conservative 0; Mismatches 671; Indels 11; Gaps 4;

QY 471 atcctaaagtgttgacatttaccctgtcttttccagaagaaatcagaacagaccac 530
DB 14613 ACCATAAGTCCTTTTAAATCAATATCTCTGCTTTCACAATAAATAATATATG 14554

QY 531 agcaaatattatgtatttcttataaaagaaagaaagaaattcatttacttatagaatt 590
DB 14553 ACCTGTAAATAATATATTAATTAATAAGAACGCAAGCAATTTTATCTGTCAGAT 14494

QY 591 aaaaagctaaagcgtgaaacggtgaagcagaatttcttaaaaaataagtaaaactctac 650
DB 14493 AAAAAAAGTGTGTTCATTTTCATCAACACACATTTGTAGACATTTGACAAATTTCCACATAT 14434

QY 651 aaacttattatgtatataacatattctataaagaaactcaaatatataatgataaactcatt 710
DB 14433 AGAATTAATATATATATCAAAACATAAATCTTCAATGGAGTATATAATATATTTT 14374

QY 711 ttaacaaatttctatgaattataaaaaaagtcacttttgacacttaaaaggttga 770
DB 14373 TTTACAAATAAATATATCTACTATATCTGTATATGTTCAATATAAGCGGAGAACTTTGCA 14314

QY 771 caataaccgtctctccaaaaaatacaaaacatttataatttctaaactatggtgtgaa 830
DB 14313 GAAAAATAAAC-----AACAGAGTTTAAATAATGTACTCAAAACAAATAAATAATCA 14260

QY 831 ttttgcgaaatacaaaagaaagaggttctctatataagtttctatttctattgtatca 890
DB 14259 TAGTATATAATTAATCAAGAAAAAATAATCCCTCACAATATTAGAACCCCTATTCCTTGT 14200

QY 891 aacttcaaatctcatgtaatttgaagaaagaaataagataatgttgtttttgtt 950
DB 14199 TTCAGAAACAGAGATAATGTTTCCCAATGATCCCAATTTGAACAAAGATCTGGACTTATT 14140

QY 951 tcttattgtacatttctatggaataatattcatacaaaaaataatgtattttaaataatg 1010
DB 14139 ATTATTATTATAAGTCTGTTTATTACATCGGAAGATTTCTTAAAGAGATTAATTAATCA 14080

QY 1011 agagattaccatcaaaaggtcgaacttatataaaaaaggttaataactaaacaatacat 1070
DB 14079 ACTGCCATAAATGAATTAAT-TACTTTGAAGGCCAAATTAACAAGAAATAAATAATATAT 14021

QY 1071 gtgatcacatcaatgacagttttgtatcttaaaatagaatgattgagcaaacctcaaaa 1130
DB 14020 ATATATACACATCTTCACAGTTTGTTCATTCATATAATATATACTAAATACAAATTCATTC 13961

QY 1131 atgtcttcttagatcaaaaaatcttcttcttagcttatttaaagccggagtgctcaactct 1190
DB 13960 ACTTTTGAATTTATACATCTATCATGTTTAGGGACAAACGACAATTTATCATATATGTTACTTTT 13901

QY 1191 cctcctctgtagaacttttttcaaatcttttttcttcttcttcttcttcttcttcttctt 1249
DB 13900 CTATGTTACAAGATTAAATGTCATTTAATTTTGTCTTAAACATGTTACGTGGAAAG 13841

QY 1250 taatgggcataatatttgggtttttaaataagtcctatagatttttttaggcacatctctaatc 1309
DB 13840 TTAAGCTTAAATATATTGTCAAAAAAGGAAAGAGATTATTCTTTTAAACCAGATTAAAAA 13781

QY 1310 acgacaaaatctcctaaattgtacacattttaaacttaaagattatgcattcaacaactct 1369
DB 13780 TGAATTAAGGACATCTCTTTTAAATAAGAAATAATATATATATATATATATATATATAT 13721

QY 1370 taaaatatatatatatatatatatatatatatatatatatatatatatatatatatata 1429
DB 13720 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 13661

QY 1430 acgataactccttactcacaattagcccaaaaaaacatcccaatgcattttaaactag 1489
DB 13660 GAGGATG---CGCTTCACACGCTGTGTTTCATGTGTAATTTTATAGATACGTTAGAAATGAC 13604

QY 1490 gaatttcaaaaactcaaatagggttggttggttaaaaaaaaacaaataagtagatgtagacata 1549
DB 13603 CAAATCTCATGGAATATATATATGTAATCGTAATGAATAAATAAATGTAACAAATTTACAA 13544

QY 1550 cgtaccttttaaaataatactactcatcgaagctttttaaatttgcgaataaatacat 1609
DB 13543 ATTCAAGACTTACGGCAATATAAACATTTCTAGAAAACTTAATTAAGAGTCTTAACAGAG 13484

QY 1610 ttactctcaaatcaaatcaatttaataatgaatgcataattctgttaattcttcttcttctt 1669
DB 13483 TGAATAAAAAATTTAAAGAAACATAATATTAAAGAAAAAATAAATAATATAGACATTATTC 13424

QY 1670 ccata 1674
DB 13423 TTATA 13419

RESULT 11
US-08-998-416-288/C
; Sequence 288, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPYII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

	Query Match	2.0%	Score	67.6;	DB 4;	Length	615;			
	Best Local Similarity	47.7%;	Pred. No.	0.00032;						
	Matches	232;	Conservative	0;	Mismatches	249;	Indels	5;	Gaps	1;
Qy	1218	aatctttcttccaaaaatcaataatgagtcgaatggcgataaatcttgggttttaatta	1277							
Db	524	AAATATATTTTATTAATATTAATAAATTTAATATAGAAAATTTAAAGCTTAAAAATTAATTTT	465							
Qy	1278	agtcacatagattttttagggaccatctctacacgcagcaaatatctctaaattgttaacacat	1337							
Db	464	TAATAATAATCTTTATAAAAAGTTAAATAATAATAAAATCAACATAAATTTTATAAAAATAG	405							
Qy	1338	ttaaaacctaaagatattgcattccacaatccttaaaaaatatatatatatatatatat	1397							
Db	404	ATATATTAATAAANAATTAATTTTCAACATTTTAATAAATAATTAATCTTTTATTAATAATAA	345							
Qy	1398	atatatatatatatgaaagtatatagaaacgataaactccttactcaacaattagc	1457							
Db	344	AAATATATTTTAAATAAACAAATTAATAAATAATATATATTAATTTATTCATAATCTATT--	286							
Qy	1458	ccaaaaaaacatccaatatgycattttaaactaggaaattttaacaaaactcaaatagggttgg	1517							
Db	287	--TAATAATTTTATTAAGAAAAATAATAATATCTAATAATAATTTTAATACTAAATTTAAAA	230							
Qy	1518	agttcaaaaaaaacaaatagtagatgfcacatagtcaccttcaaaaaatatatactcatc	1577							
Db	229	TTTGAAACATAGACATAATAGTATTTCATATTAATAATATTTTATTATTAATTAATAAATATT	170							
Qy	1578	gaaagttttaaatttttgcgaaattaaataacatttatctatcaattaaaaatacacatttaata	1637							
Db	169	AATAATGATGAATTAAGTAAATATATATAAATAATTAATAATAATAAGTATTTAATAATCAATA	110							

Query Match	1.98;	Score 66.2;	DB 6;	Length 3095;
Best Local Similarity	44.3%;	Prod. No. 0.00069;		
Matches 356;	Conservative 0;	Mismatches 443;	Indels 4;	Gaps 2;
Qy 632	aaaatagtaaaactgtacaaaacttattt	gtgtatataa	catatc	tataaagaactc 691
Db 2165	aaaggattcatcaaaataaaaataaaa	agaatcatct	ttttataca	tatatctcaaa 2224
Qy 692	aaatatgataaatcatcttttaacaaa	atttctcatga	attataata	aaaaaagtcactt 751
Db 2225	agaaatttaaaaagtatc	caaaactatg	taagtg	tattgaatgaatgatggtg 2284
Qy 752	ttgacattaaaaaggttgacaata	accogtctct	cccaaaaa	aaaaaacatttataat 811
Db 2285	tattcaagttgtaagtaacaatt	taaggaatag	caagga	tatagtaatttttccaaa 2344
Qy 812	ttctaaaactatggtgtaattttgct	gaatcaaaa	agaagaag	gaatttcttatataat 871
Db 2345	acattttaaataattcaacaaaa	aaaaaa	aaaaaa	aaaaaaatttcaataaaat 2404
Qy 872	aagtttcattattgtatcaaac	ttccaattt	ccatgaatt	tgaaaggaagaaaaaattaaag 931
Db 2405	tttttcttattata	tgtaactaat	tattattata	aaataatatataataaaa 2464
Qy 932	ataaaagtgttttgttttctt	atgtttacatt	ttctgga	ataataattcatacaaaaa 991
Db 2465	gaatactaggatttctgtatata	tatagaaa	taaaattc	attgtattattatagaaaaaaa 2524
Qy 992	aatgtattttaatatgatgatgag	agattaccat	cccaaa	agggtcgaaacttataataaaacagt 1051
Db 2525	taaaaataaaataattttctt	cttagttg	tattata	tattctcaaaaataataataaaa 2584
Qy 1052	ta---aaactaaacaatacat	atgataatg	atacaaat	caatgacaggttttgatcttaaaaaataga 1108
Db 2585	tatgaacaatttcaagtata	tattttaa	atctgaa	aggtatagatcgtataaaaaaatttt 2644
Qy 1109	aatgatgtgacaaacctc	caaaaa	tgctctct	taggatacaaaaaatcttctcttagctta 1168
Db 2645	ctggctttataataataaaa	taagaaaatt	tattatt	atgattattttaagaataatta 2704
Qy 1169	ttaagcgggagttcaactct	ctctcctt	gtgagact	ttttgttttttcaaatctttttct 1228
Db 2705	tataataatgataataaat	gaaagaaa	aaaaaa	aaaaatttttttattatcatctttttca 2764
Qy 1229	ttcaaaaaatacaataant	tagttaatg	ggcata	ataattattgtgttttaattaaagtcacatagat 1288
Db 2765	taaatattataaaa	caaatatt	tatatata	aaaaaacctgggataaaataattttagtataatttat 2824

```

;
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: Rh 32
; US-08-883-795A-36

Query Match      1.98; Score 64.8; DB 2; Length 665;
Best Local Similarity 49.6%; Pred. No. 0.001;
Matches 257; Conservative 0; Mismatches 247; Indels 14; Gaps 3;

Qy 518 ATAAAAATGTAATATAAACTTTAATATAAAATGTAATATAAACTTTATAA 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1328 tgaacacatttaaaacttaaaagcattgcacacaccccttaaaatatatatat 1387
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 458 AATATGTAATATAAAATATGTAATATAAAACATTTTAAATATAAAATATGTAATATAA 399
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1388 atatatatatatatatatataatgaagtttatagaaacgataaacctctactc 1447
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 398 ACATTTTAAATATAAAATATGTAATATAAAACATTTTAAATATAAAATATGTAATATAA 339
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1448 acaatttgcgcacaaacacatccata-atgcatttaaacctagggaattttacaaactca 1506
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 338 ACATTTTAAATATAAAATATGTAATATAAAACATTTTAAATATAAAATATGTAATATAA 279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1507 aataggttggttagttaaataaaacacaaatagtagatgacatacgtaacctttataaaatat 1566
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 278 ACATTTTAAATATAAAATATGTAATATAAAACATTTTAAATATAAAATATGTAATATAA 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1567 atactcatatcgaaagtttttaaaatttgcgaatttaaacattttatctatcaattaa 1626
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 218 ATATTTTAAATATAAAATATGTAATATAAAATATTTTAAATATAAAATATTTAATATAA 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1627 tacatttaaatgcatatctctgtaatatctatctatctatctatctatctatctatctat 1686
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158 ATATTTTAAATATAAAATATGTAATATAAAATATTTTAAATATAAAATATTTAATATAA 99
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1687 aaaaataacatcaataatagttttaaacttaacaaacgcttagggaaagttgaccttaac 1746
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 98 ATATTTTAAATATAAAATATGTAATATAAAATATTTTAAATATAAAATATTTAATATAA 39
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1747 tagcttgattga 1758
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38 AATATTTAATTA 27
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 15

```

US-08-883-795A-36
; Sequence 36, Application US/08883795A
; Patent No. 5985607
; GENERAL INFORMATION:
; APPLICANT: Delcuve, Genevieve
; APPLICANT: Awang, Gregor
; TITLE OF INVENTION: Recombinant DNA Molecules and Expression
; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/883,795A
; FILING DATE: 27-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 7841-062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 665 base pairs
; TYPE: nucleic acid
```

```

Search completed: July 13, 2001, 00:22:31
Job time: 8179 sec
```





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2001, 22:21:17 ; Search time 218.86 Seconds

(without alignments)  
9823.335 Million cell updates/sec

Title: PCT-US01-13059-3

Perfect score: 3424

Sequence: 1 aagctaaagctaccacat.....aacggcatagctgactcagt 3424

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_0601.\*

- 1: /SIDS1/gcgdata/geneseq/geneseq/NA1980.DAT.\*
- 2: /SIDS1/gcgdata/geneseq/geneseq/NA1981.DAT.\*
- 3: /SIDS1/gcgdata/geneseq/geneseq/NA1982.DAT.\*
- 4: /SIDS1/gcgdata/geneseq/geneseq/NA1983.DAT.\*
- 5: /SIDS1/gcgdata/geneseq/geneseq/NA1984.DAT.\*
- 6: /SIDS1/gcgdata/geneseq/geneseq/NA1985.DAT.\*
- 7: /SIDS1/gcgdata/geneseq/geneseq/NA1986.DAT.\*
- 8: /SIDS1/gcgdata/geneseq/geneseq/NA1987.DAT.\*
- 9: /SIDS1/gcgdata/geneseq/geneseq/NA1988.DAT.\*
- 10: /SIDS1/gcgdata/geneseq/geneseq/NA1989.DAT.\*
- 11: /SIDS1/gcgdata/geneseq/geneseq/NA1990.DAT.\*
- 12: /SIDS1/gcgdata/geneseq/geneseq/NA1991.DAT.\*
- 13: /SIDS1/gcgdata/geneseq/geneseq/NA1992.DAT.\*
- 14: /SIDS1/gcgdata/geneseq/geneseq/NA1993.DAT.\*
- 15: /SIDS1/gcgdata/geneseq/geneseq/NA1994.DAT.\*
- 16: /SIDS1/gcgdata/geneseq/geneseq/NA1995.DAT.\*
- 17: /SIDS1/gcgdata/geneseq/geneseq/NA1996.DAT.\*
- 18: /SIDS1/gcgdata/geneseq/geneseq/NA1997.DAT.\*
- 19: /SIDS1/gcgdata/geneseq/geneseq/NA1998.DAT.\*
- 20: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT.\*
- 21: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT.\*
- 22: /SIDS1/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	259	7.6	936	22 AAF58252	Oligonucleotide D1
C 2	259	7.6	936	22 AAF58254	Oligonucleotide D1
C 3	259	7.6	936	22 AAF58257	Oligonucleotide D1
C 4	259	7.6	936	22 AAF58259	Oligonucleotide D2
C 5	259	7.6	936	22 AAF58262	Oligonucleotide D2
C 6	259	7.6	936	22 AAF58255	Oligonucleotide D1
C 7	252	7.4	936	22 AAF58252	Oligonucleotide D1
C 8	252	7.4	936	22 AAF58254	Oligonucleotide D1
C 9	252	7.4	936	22 AAF58257	Oligonucleotide D1
C 10	252	7.4	936	22 AAF58259	Oligonucleotide D1
C 11	252	7.4	936	22 AAF58262	Oligonucleotide D2

12	252	7.4	938	22 AAF58255	Oligonucleotide D1
C 13	115.6	3.4	19124	18 AAT72882	Plasmodium var-7 p
C 14	115.6	3.4	19124	21 AAZ98287	Plasmodium var-7 p
15	96.4	2.8	3549	21 AAA70223	Plasmodium falciparum
16	95.2	2.8	8310	20 AAZ29911	cDNA encoding a SC
17	83.4	2.4	6644	20 AAX33181	Base sequence of t
18	83.4	2.4	7372	20 AAX33182	Base sequence of t
19	83.4	2.4	7797	20 AAX33180	Cowpox virus bsr f
20	83.4	2.4	7996	20 AAX33184	Base sequence of t
21	83.4	2.4	19124	18 AAT72882	Plasmodium var-7 p
22	83.4	2.4	19124	21 AAZ98287	Plasmodium var-7 p
C 23	81.2	2.4	4590	7 AAN60472	Sequence encoding
24	80.2	2.3	244	22 AAF58238	Oligonucleotide D1
C 25	79.6	2.3	244	22 AAF58238	Oligonucleotide D1
26	79.4	2.3	8920	15 AAQ62924	Carbamoyl-phosphat
27	79	2.3	6124	11 AAQ03568	Sequence encoding
C 28	78.2	2.3	1864	8 AAN71405	Sequence of ANS-1
29	77.2	2.3	6888	21 AAA70114	Plasmodium falciparum
30	76.4	2.2	1864	8 AAN71405	Sequence of ANS-1
31	76	2.2	15577	19 AAV35616	SHOX gene prelimin
32	75.6	2.2	32367	19 AAV35620	Human SHOX (short
33	75.6	2.2	1671	13 AAQ24134	50 kD subunit of S
C 34	75.2	2.2	8310	20 AAZ29911	cDNA encoding a SC
35	74.8	2.2	2503	15 AAO53480	PNPX30 xylanase CD
36	74.8	2.2	15418	21 AAA63785	Nucleotide sequenc
C 37	74.8	2.2	15418	21 AAA63785	Nucleotide sequenc
38	74.6	2.2	1230	10 AAN90223	Malaria-specific D
39	74.4	2.2	7620	21 AAA70132	Plasmodium falcipa
C 40	72.4	2.1	5852	12 AAQ11710	Dictyostelium plas
41	72.4	2.1	9789	17 AAT41852	cDNA encoding Plas
C 42	72.2	2.1	1864	15 AAQ78892	Aspergillus nidula
43	71.8	2.1	6152	18 AAT78867	P. falciparum live
44	71.4	2.1	1431	21 AAZ37082	DNA sequence encod
C 45	71.4	2.1	2104	13 AAQ25273	DNA sequence encod

ALIGNMENTS

RESULT 1	
AAF58252/c	
ID AAF58252 standard; DNA; 936 BP.	
XX	
AC AAF58252;	
XX	
DT 24-APR-2001 (first entry)	
XX	
DE Oligonucleotide D1835.	
XX	
KW Electron-transfer group; ETW; mismatch; genotyping;	
KW gene expression; ss.	
XX	
OS Synthetic.	
XX	
PN WO200107665-A2.	
XX	
PD 01-FEB-2001.	
XX	
PF 26-JUL-2000; 2000WO-US20476.	
XX	
PR 26-JUL-1999; 99US-0145695.	
PR 17-MAR-2000; 2000US-0190259.	
XX	
(CLIN-) CLINICAL MICRO SENSORS INC.	
Umek RM;	
XX	
DR WPI; 2001-159728/16.	
XX	
PT Nucleic acids containing electron-transfer group, useful as labels in	
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on	
PT a single surface	
XX	

PS Example 6; Page 127; 159pp; English.

CC The present invention relates to a composition comprising two nucleic  
 CC acids each containing an electron-transfer group (ETM) having  
 CC different redox potentials. The invention is used for electronic  
 CC detection of nucleic acids, especially of substitutions (mismatches)  
 CC and single-nucleotide polymorphisms, e.g. for genotyping,  
 CC monitoring gene expression.  
 QQ - Sequence 936 BP: 4 A: 139 C: 10 G: 7 T: 776 other:  
 SQ

Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

```
Query Match          7.6%; Score 259; DB 22; Length 936;
Best Local Similarity 0.9%; Pred. NO. 4.5e-30;
Matches 7; Conservative 600; Mismatches 180; Indels 0; Gaps 0;
```

[illegible]

## RESULT 2

AAF58254/c  
ID AAF58254 standard; DNA; 936 BP.

AA  
AC  
AAF58254;

DT 24-APR-2001 (first entry)

DE Oligonucleotide D1875.

Electron-transfer group; ETM; mismatch; genotyping;  
KW  
KW  
gene expression; ss.

OS Synthetic.

PN WO200107665-A2.

PD 01-FEB-2001.

26-JUL-2000; 2000WO-US20476.

PR 26-JUL-1999; 99US-0145695

[illegible]XX  
XX  
/

XX

**XX**

PT hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface

PS Example 6; Page 127; 159pp; English.

CC The present invention relates to a composition comprising two nucleic  
CC acids each containing an electron-transfer group (ETM) having  
CC different redox potentials. The invention is used for electronic  
CC detection of nucleic acids, especially of substitutions (mismatches)  
CC and single-nucleotide polymorphisms, e.g. for genotyping,  
CC monitoring gene expression.

SQ Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

Query Match 7.68; Score 259; DB 22; Length 936;

Best Local Similarity	0.9%	Pled. NO.	4.5e-30,
Matches	7;	Conservative	600;
		Mismatches	180;
		Indels	0;
		Gaps	0;

Qy 625 ttctaaaaaaataqtaaaactqctacaaactatttatqqtatataacatatctataa 684

Db 787 WWWWWW...WWW 728

Qy 685 gaaactcaaatatgataaatcattttaacaaaaatttctatgaaattataataaaaaa 744

**Db** 727 WWWWWW...WWWGWWWWW...WWW 668

Qy 745 gtcaactttgacacttaaaaggttgacaataaccgtctctccaaaaaaatcaaaacat 804

Db 667 WWWWWW...WWW 608

QY 805 ttataatttctaaaactatggtgttaattttgctgaaatcaaaaqaaaggatttct 864

[illegible]

Qy 865 atatcaagttcattattgtatatcaaaactttcaaatttcattgtaatttgaaaggaaaa 924

**D<sub>b</sub>**     547 WWWWWWTTTTTTTTTTTTTTTTTTTTTTTWWWWWWWWWGTTTTWW  
488

Qy 925 aattaagataataatgtgtgtttttgtttcttatgtttacattttcatggaatatattcat 984

[illegible]

CC The present invention relates to a composition comprising two nucleic  
CC acids each containing an electron-transfer group (ETM) having  
CC different redox potentials. The invention is used for electronic  
CC

[illegible]

[illegible]

KW	Electron-transfer group; ETW; mismatch; genotyping;
KW	gene expression; ss.
XX	
OS	Synthetic.
XX	
PN	WO200107665-A2.
XX	
PD	01-FEB-2001.
XX	
PF	26-JUL-2000; 2000WO-US20476.
XX	
PR	26-JUL-1999; 99US-0145695.
PR	17-MAR-2000; 2000US-0190259.
XX	
PA	(CLIN-) CLINICAL MICRO SENSORS INC.
XX	
PI	Umek RM;
XX	
DR	WPI; 2001-159728/16.
XX	
PT	Nucleic acids containing electron-transfer group, useful as labels in
PT	hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT	a single surface
XX	
PS	Example 6; Page 127; 159pp; English.
XX	
CC	The present invention relates to a composition comprising two nucleic
CC	acids each containing an electron-transfer group (ETW) having
CC	different redox potentials. The invention is used for electronic
CC	detection of nucleic acids, especially of substitutions (mismatches)
CC	and single-nucleotide polymorphisms, e.g. for genotyping,
CC	monitoring gene expression.
XX	
SQ	Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;
<p>Query Match 7.6%; Score 259; DB 22; Length 938;            Best Local Similarity 0.9%; Pred. No. 4.5e-30;            Matches 7; Conservative 600; Mismatches 180; Indels 0; Gaps</p>	
Qy	625 ttctaaaaaaatagtaaaactgctcaaaacttatttattgtgtatataacacatatcataaa 684
Db	787 WWW 728
Qy	685 gaactcaaatatgatataaatttatacaaaaatttctatgaattataataaaaaa 744
Db	727 WWW 668
Qy	745 gtcaattttgacacttaaaagggttgacataaccgtctctccaaaaaaaaatacaaacat 804
Db	667 WWW 608
Qy	805 ttataattctcaaacatggttgtaatttgcgtgaatacaaaaagagaagaagatttct 864
Db	607 WWW 548
Qy	865 atatcataagtttcaattgtatcaaaactttcaaatctcatgtaatttgaagagaaaa 924
Db	547 WWW 488
Qy	925 aattaagataaagtgtgttttgtttcttattgttacatttcatggaatatattcat 984
Db	487 WWW 428
Qy	985 acaaaaaaatgtattttaatatgatgagagattacacatccaaaaagggtcgaaactataaa 1044
Db	427 WWW 368
Qy	1045 acaagtttaactaaacaatacatatggtgatcacatcaatcaatgatttgatcttaaaa 1104
Db	367 WWW 308
Qy	1105 tagaaatgattgagcaaacctcaaaaatgtcttcttaggatcacaaaaatcttctcttag 1164





[illegible]

Qy	690	tcaaatatagataaatacattttaacacaaattctctatgaattataataaaaaaagtcac	749
Db	61	#####	120
Qy	750	ttttgacacttaaaaggttgacataaacgctctctccaaaaaaatacaaacatttata	809
Db	121	#####	180
Qy	810	attctaaactatgggtgtaatttttctgtaaatcaaaagaaaaagaagattctctatc	869
Db	181	#####	240
Qy	870	ataagtttcattatgtatacaaaccttcaaatattcatgtaatttgaaaggaaaaaatta	929
Db	241	#####	300
Qy	930	agataaagtgtttttgttcttatgttcaattttcatgaatatatttcataacaa	989
Db	301	#####	360
Qy	990	aaatgtattttaatgatgatgagattaccatccaaagtcgaacctataataacaa	1049
Db	361	#####	420
Qy	1050	gttaataactaaacaatacatgtatcaacaatcaatgacagtttggatcttaaatagaa	1109
Db	421	#####	480
Qy	1110	atgattgacaaactcaaaaatgtctcttaggatcacaaaatcttctttagcttat	1169
Db	481	#####	540
Qy	1170	taaagcggaggtcgaactctctccctgtgaccttttgtttcgaactctttctt	1229
Db	541	#####	600
Qy	1230	tcaaaaatacaataatagtttaatgggcataataattggttttaattaaagtcocatagatt	1289
Db	601	#####	660
Qy	1290	ttttaaggaccatctctaactcagcagacaaatatactctaattgttaacacatttaaaactaaa	1349
Db	661	#####	720
Qy	1350	agattgcatccaaactctaaaatatatatatatatatatatatatatatatat	1409
Db	721	#####	780
Qy	1410	atatat 1415	
Db	781	#####	

RESULT	10
--------	----

AAF58259  
ID AAF58259 standard; DNA; 936 BP.

XX AAF58259;

XX  
DT 24-APR-2001 (first entry)

XX  
DE Oligonucleotide D2004.

xx Electron-transfer group; ETM; mismatch; genotyping;  
kw gene expression; ss.

XX Synthetic.

XX  
PN WO200107665-A2.

XX PD 01-FEB-2001.

XX  
PF 26-JUL-2000; 2000WO-US20476.

26-JUL-1999; 990US-0145695.  
17-MAR-2000; 2000US-0190259.  
(CLIN-) CLINICAL MICRO SENSORS INC.  
Umek RM;  
WPI; 2001-159728/16.  
Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface  
Example 6; Page 128; 159pp; English.  
The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.  
Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;

Query Match 7.4%; Score 252; DB 22; Length 936;  
Best Local Similarity 0.4%; Pred. No. 4.8e-29;  
Matches 3; Conservative 599; Mismatches 184; Indels

Qy	630	aaaaaatagtgtaaacctgctacaaacttat	tttatgtgtatataacacatactctataagaac	689
Db	1	#####	#####	60
Qy	690	tcaaatatgataaatcattttaacaaat	tctatgaaattataataaaaaagtccac	749
Db	61	#####	#####	120
Qy	750	ttttgacactaaaggttgacaaataacog	tctctccaaaaaaaatccaaacatttata	809
Db	121	#####	#####	180
Qy	810	atttctaaaaactggtgtaattttgtcgt	gaaatcaaaaagaaagagatttctatc	869
Db	181	#####	#####	240
Qy	870	ataagttcattatgtatcaaaacttccaa	atttcattgttaattgaaaggaaaaatta	929
Db	241	#####	#####	300
Qy	930	agataaatgtgtgtttgttcttattgt	tacattttcatggaaatataatcatacaa	989
Db	301	#####	#####	360
Qy	990	aaaatgtatttaatatgatgagattacc	atccatccaaaaggtcgaaacttataaaa	1049
Db	361	#####	#####	420
Qy	1050	gttaataactaaacaatacatgtgatc	acaatccaatgacagtttgttgccttaaa	1109
Db	421	#####	#####	480
Qy	1110	atgattgacgaacacctcaaaaatgt	tctctttaggatacacaatacttctctt	1169
Db	481	#####	#####	540
Qy	1170	taaacgcgagggttcaactctctccct	tgttagacttttgtttcgaactcttttctt	1229
Db	541	#####	#####	600
Qy	1230	tcaaaaaatcaataatagtttaatgg	cataataatttggtttttaagtcctatatt	1289
Db	601	#####	#####	660

[illegible][illegible]

QY 1410 actual 1410  
: : : : :  
Db 781 WWWWW 786

RESULT 12  
AAF58255  
ID AAF58255 standard; DNA; 938 BP.  
XX  
XX AC AAF58255;  
XX  
XX  
XX 24-APR-2001 (first entry)  
XX  
XX DE  
XX Oligonucleotide D1876.  
XX

gene expression; ss.  
XX  
XX  
OS  
XX  
XX  
PN  
XX  
XX  
PD  
XX  
XX  
PF  
XX  
XX  
PR  
PR  
PR  
XX  
PA  
(CLIN-) CLINICAL MICRO SENSORS INC.

```
XX Umek RM;
PI
XX
DR WPI: 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface
XX
XX
PS Example 6; Page 127; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;

Query Match 7.4%; Score 252; DB 22; Length 938;
Best Local Similarity 0.4%; Pred. No. 4.8e-29;
Matches 3; Conservative 599; Mismatches 184; Indels 0; Gaps 0;

Qy 630 aaaaaatagtagtaaaactgtatctatttattgtgtatataacatctatataaagaac 689
Db 1 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
Qy 690 tcaaatatgatataaacttttaacaaaatttctatgaattataataaaaaaagtcac 749
Db 61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 120
Qy 750 ttitgacacttaaaagttgacataaacctctctcacaataaaacatttata 809
Db 121 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 180
Qy 810 attctaaactatggtgtatatttctgtaaaacaaaagaaggatttctatc 869
Db 181 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 240
Qy 870 ataagtttcattatgtatcaaaacttcaaaattcattgaatttgaaaggaataa 929
Db 241 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 300
Qy 930 agataatgtgtttttgtttcttattgtacatttctcgaataatattcatacaa 989
Db 301 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 360
Qy 990 aaaaatgtatttaatatgatgagattacacatccaaaaggctgaacttatataa 1049
Db 361 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 420
Qy 1050 gtaataactaaacatacatgtatcacaaatcaatgacagtttgcattcaaataga 1109
Db 421 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 480
Qy 1110 atgattgacaaacctcaaaaatgtcttcttaggatcacaaaacttcttcattat 1169
Db 481 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 540
Qy 1170 taagcgggagttcaactctctcccttgtagacttttttttcaaatcttttctt 1229
Db 541 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 600
Qy 1230 tcaaaaaatcaataattagtttaaggcataaatttggttttaatttaagtcata 1289
Db 601 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 660
Qy 1290 ttttaggacctctctaatacagcaaatatccataaattgtaacacatttaaaacttaa 1349
Db 661 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 720
```

```
Qy 1350 agtattgcattcaacaccttaaaatatatatatatatatatatatatatatatat 1409
Db 721 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 780
Qy 1410 atatat 1415
Db 781 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 786

RESULT 13
AAT72882/c
ID AAT72882 standard; cDNA; 19124 BP.
XX
AC AAT72882;
DT 12-SEP-1997 (first entry)
XX
DE Plasmodium var-7 gene.
XX
KW DBL gene family; SABP; sialic acid binding protein; vaccine; therapy;
KW Duffy binding like gene; Duffy antigen binding protein; erythrocyte;
KW DABP; merozoite; malaria; var-1; var-2; var-3; var-7; immune response;
KW Plasmodium; ss.
XX
OS Plasmodium vivax.
XX
PH Plasmodium falciparum.
XX
FT Key Location/Qualifiers
FT exon 7317..15139
FT /tag= a
FT /number= 1
FT Intron 15140..16205
FT /tag= b
FT /number= 1
FT exon 16206..17552
FT /tag= c
FT /number= 2
FT /note= "no stop codon given"
XX
WO9640766-A2.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US09508.
XX
PR 07-JUN-1995; 95US-0487826.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Chitnis C, Miller LH, Peterson DS, Sim KL, Su X;
PI Wellens TE;
XX
DR WPI: 1997-052231/05.
DR P-PSDB; AAW22475.
XX
XX New malaria vaccines - contains cysteine-rich DBL family protein
XX binding domains homologous domains of the Duffy and sialic acid
XX binding proteins
XX
PS Claim 4; Page 56-61; 96pp; English.
XX
CC This sequence represents the var-7 gene of Plasmodium. Var-7 belongs to
CC the Duffy binding like (DBL) family of genes which have homology to the
CC Duffy antigen binding protein (DABP) and sialic acid binding protein
CC (SABP) conserved regions (see AAT72889 and AAT72888 respectively). The
CC var family of genes modulate cytoadherence and antigenic variation of
CC Plasmodium infected erythrocytes. SABP and the Duffy antigen binding
CC protein (DABP) are soluble proteins that appear in the culture
CC supernatant after infected erythrocytes release merozoites. DABP and
CC SABP mediate the binding of merozoites and schizonts to the erythrocyte
CC surface. These proteins are necessary for erythrocyte invasion by the
CC parasite. This sequence can be used in the compositions of the
CC invention. The compositions are for the treatment and prevention of
```

CC malaria, and comprise either a nucleotide sequence or encoded polypeptide  
CC of the var-1, var-2, var-3 or var-7 genes of the DBL gene family, a  
CC family of genes having homology with conserved regions of DABP and SAPP.  
CC The compositions are used for the treatment and prevention of malaria.  
CC They are also used in the preparation of vaccines for inducing a  
CC protective immune response in a mammal to Plasmodium merozoites  
CC (especially Plasmodium falciparum or Plasmodium vivax).  
XX  
SQ Sequence 19124 BP; 7824 A; 2190 C; 2790 G; 6320 T; 0 other;

Query Match		3.4%;	Score 115.6;	DB 18;	Length 19124;
Best Local Similarity		43.3%;	Pred. No. 4.1e-09;		
Matches 917;		Conservative 0;	Mismatches 1174;	Indels 27;	Gaps 7;
Qy	390	tatattgctctttcttactacttcttttcgagactatttgccttatttttgcctcaaac	449		
Db	7083	TATATATATTAATAGTGTGACTATTATATATTAATTAACAATATGATATCTTAATAAATA	7024		
Qy	450	tttttggattgcacttttcgactcctaaagtggttgacaatttacctgcttttctcca	509		
Db	7023	ATATTTAAATGATATATAATAATAATTAATTAAGTTTATATATATTTTAAAAAAATATATA	6964		
Qy	510	agaaaacacagacagacacagcaaatattatgtatttcttattaaagaaagaaaga	569		
Db	6963	TATATATTAATAAATTTAATAAATTTAATAATTTCTACAAAAAATAATAATCAGA	6904		
Qy	570	attcatattacttagaattaaagcagtggaacgtgaaacgtgaaagcagaatttcta	629		
Db	6903	AATATTATATTTTATGATTCCTTTTATTTATCTATTTAAATTAATATATATATTTT	6844		
Qy	630	aaaaaataagtaaaactgctacaaacttatttggtgtatataacatctctataaagaac	689		
Db	6843	TATGTTTTTATTTAAAGTAAATTAATAATGAG-----AAAAAATAACGAAATAC	6789		
Qy	690	tcaatatatgataatcattttaacaaattctcatgaaattataataaaaaagtcac	749		
Db	6788	AAACATATAAAGAGTATATATGCAACGTGTTTATATATTTAATTTAATTAACATTATATA	6729		
Qy	750	ttttgacacttaaaaggttgacaataaacgcgtctctccaaaaaaatacaaacattata	809		
Db	6728	TGTATATTTTTTTTGCATTTATTTTAAATTTATATATATATATATATATATAGATAACA	6669		
Qy	810	attctcaaacactagtgtaatttctgctgaaatcaaaagaaagaaagatttctcatc	869		
Db	6668	AAAGAGACAAATATGTTCTGCTTCTCTCTATCTATATATTTATCTCATATTTATATAT	6609		
Qy	870	ataagtttcattatgtatcaaaacttcaaaatttcatgtaatttgaaaggaagaaatta	929		
Db	6608	CATATATATAAATTAATAGATACATATTTCTTGTATGTTGTTATATATAATAAGTA	6549		
Qy	930	agata-taagttgttttcttcttcttcttcttcttcttcttcttcttcttcttcttctt	988		
Db	6548	GTAT	6489		
Qy	989	aaaaatgtattttaaataatgatgatgagattaccatccaaaaggtcgcaacttatataa	1048		
Db	6488	ATAGCATCAAAAAAATAATGATAAATAAATAACAGGAAAAATATATATATATATAT	6429		
Qy	1049	agttaactaaacatacatatgtatcacaatcaatgacagtttttgatctttaaataaga	1108		
Db	6428	ATATTATATAAATAAATGTTTTATCATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	6369		
Qy	1109	aatgattgagcaaacctcaaaaatgcttcttcttcttcttcttcttcttcttcttctt	1166		
Db	6368	TGCATTTTATGATTTCAAAATTTTATGCTATATAATAATAAATAAAGTAAATAACACA	6309		
Qy	1167	tattaagccggagtgcaact	1226		
Db	6308	TTATAAATATATATATCAAAATGAGTTATTAATAAATAAATGTTGTTGTTGTTGTTG	6249		
Qy	1227	ctttcaaaaaataaataatagtttaattgggcaataatttgggttttaataagtcacatg	1286		

Db	6248	TATAAATGAAATATATTTGTTATATATAATAATACATATATATGCTACTATATATAATA	6189		
Qy	1287	attttttaggaccatctctaaatcacgacaaatctcctaaattgtcaacattttaaactt	1346		
Db	6188	ATATCTTTAAAGTATATACTA-----AAATATATAAAATGCAATGATAAAATAGTA	6136		
Qy	1347	aaaaatgtgattcacaatcctttaaataatataatataatataatataatataatata	1406		
Db	6135	TAAATCATACATA	6076		
Qy	1407	Latatataaaggttatagaaacgataaactcctctcaacaattagcccaaaaaa	1466		
Db	6075	TGCATATATAGTAATAATTTTATATATTTCCAAATCTGATATGTTTATATATTTGTT	6016		
Qy	1467	catccataatgcattttaaacttaggaattttaaacaactcacaataggttggttagt	1526		
Db	6015	ATATTATAATAACAAAAAGACGACAAAGAGTAACAAAAAATAATAATAATAA	5956		
Qy	1527	aaaaaataagtagtgatcacatgcactttaaataatataatataatataatataatata	1586		
Db	5955	GGATGATATAATAAACAAATTTACCTATTAGTATAGTTTAAATGTC---TTTATATATAT	5899		
Qy	1587	aaatttgcgaaatcaaatcacatttctctcaatttcaataacatttcaataatgcataat	1646		
Db	5898	ATATATATATATATTTTATACAAATATTTCTTCTGAATATGCATATGTTAGTTATCTA	5839		
Qy	1647	tctgtaatatctcttcttcaatttccatagaaaccccaaaaaataaacaatcatcaaatag	1706		
Db	5838	TTTTTTATATATATGATACAAAAAATAAATAAATAAATAAATAAATAAATAAATAA	5779		
Qy	1707	ttttaacttaacaaaacggttaggaaagttgacctgaactgcttgattgacgttgac	1766		
Db	5778	ATCTATCTTAAACACAGACATATACACAAA-----TAAATATATACTTATATGTAATC	5726		
Qy	1767	tgtcaatgcgaaagcgatatttccaatatatactacatgtagtattttatggaag	1826		
Db	5725	TATTCATATATAATATATAATGAACACCCCATATATACTTATAGCCATATATACACCA	5666		
Qy	1827	ttcttaaaaggtgtgagtggttcttcttcttcttcttcttcttcttcttcttcttct	1886		
Db	5665	ATCAAAACACATAAATAATATATATATATAATATATATATATATATATATATATAT	5606		
Qy	1887	caataaattttacagagatgggataaactacatactcatgattatgaaacgctcaattta	1946		
Db	5605	TTTGATGAAATACATATATTTT--TAACTAATTTATTTATTTATATATATATATATAT	5548		
Qy	1947	tttgaaacacctcctaatacacccaaatgctcactagattccaaaacgtagaccaattgt	2006		
Db	5547	TTTTCAAAAAACTAATCATCTGTTAATAATATAATAATAATAATAATAATAATAATAA	5488		
Qy	2007	atctaactcctcaaatctcaatcaaggtatttaattaccogattggaagaaagtaacoga	2066		
Db	5487	ATAATTTTAAATATATATATATATCTTAAATATTAATTTCTTATTTATTTCTATAATAC	5428		
Qy	2067	tataattatcaaaagaaagataaagtcacagattcttaactcttcttcttcttcttctt	2126		
Db	5427	GTTAGCTATTTGTAATATATTTTACATTTTAAATAATGTAATATATGACTTATTTATTCAG	5368		
Qy	2127	aacatttgacaaaaatctcaaaagatatgtaactgttttaaaatataataaactcactgagat	2186		
Db	5367	TATTTCTTATTTTATTTTATTTTATTTTCTTATTAAGATTTAAATAATATTTTACTATTAC	5308		
Qy	2187	taattcttcagactcgtgttagctataataatgctcaaggttcttcttcttcttcttctt	2246		
Db	5307	TTGTTGTAAGATATATAATAATATATATATATGAGTCCATAATTTATTAAGCTATTGGAAC	5248		
Qy	2247	aaccttaagaatgatatattttctgtaattatgatgatataaatttgctattcattgcta	2306		
Db	5247	TACAAAAACAGGATATTTATTTCTTATTTTCTTATTTTAAATAAATAATTTATTTATTA	5188		
Qy	2307	caaacattactttaaaaaatcgatttttcttcttcttcttcttcttcttcttcttctt	2366		
Db	5187	TATAGACGATTCATATATAAATTAATTTAGTCCATTAATTTCAAAAAACACAAAT	5128		





QY	1527	aaacaaaatagtagatggtacatacgttaacttttaaataatatatactcatcgcaaatgtt	1586
Db	5955	GGATGATAAATAAAACTAATTACCTTATTAGTAGTATGTTTTAATGTC--TTTATATATAT	5899
QY	1587	aaatttcgcgaattaaatacacatttatcatcaattaaatacacattaaataatgacataat	1646
Db	5898	ATATATATATATATTTATACAAAATATTATTCTTGTAATATGCATATGTGGTAGTACTA	5839
QY	1647	tctgtaatatctatctttaatttccatataagaaccaacaataaacatatacaaatag	1706
Db	5838	TTTTTTATATATGTACAAAAAANAANAANCTATAACCCTTAACTGTAAATGCCAT	5779
QY	1707	ttttaaccttaacaaaaogtgaggaaaagtggacctaaactagcttgattgacgllgaac	1766
Db	5778	ATCTATCTTAAACACAGACATATACACAAA-----TAAATATATTAACTTATATGTA	5726
QY	1767	tgtcaatgcgaaagcgatatttccaatatatactacatgtagtagtattattatataggaag	1826
Db	5725	TATTCACATATAAANTATTAATGAACAACCACCATATATACTTATAGCCTTATATACACCA	5666
QY	1827	tttctaanaaggtgttgagtggtgactgttggtggaggtactatttttctctcttg	1886
Db	5665	ATCAAANAACACATAAATATATATATATATATATATATATATATATATATTAAGTTAGTAT	5606
QY	1887	ccataaatatttacgagtagggataaactcatactcatgatattatgaacgcctcactta	1946
Db	5605	TTTGTATGAATACTATTATT--TAACTAATTTATTATTATTATATCATATTTAATATAT	5548
QY	1947	tttgaanaacctctaatacaccaaataatgtoactagattcccanaacgtagacaaatgt	2006
Db	5547	TTTTCAAAAACCTAATCATGTGTTAATATAATATAATATACATAATAATGCATAAAAAAT	5488
QY	2007	atctaactcnaattctcaatcaaatgagtagtaattaccogtggtaagaaaagttaacoga	2066
Db	5487	ATAATTTTAAATATTTATATTATTCCTTAATATTAAATTCCTATTATTCTATAATACAT	5428
QY	2067	tataattatcaaaagaagaaagaatgcacagattcttaactctcttatttttgggtatg	2126
Db	5427	GTTACGTAATTTGTAATTATGTTACATTTTAAAAAATGATAAATATGACTTATTATTTCAG	5368
QY	2127	aacatttgtacaaaaatctcaaaagatatgtaactgttttaaaataataaattcacctgagat	2186
Db	5367	TATTCCTTATTTTATTATTTTATTTTCCTTATTAAAGATTAATAATAATTTACTTTATTAC	5308
QY	2187	taattctcagactcgtgttagctataataatgtcaagagttctctgttttcaagaaa	2246
Db	5307	TTGTTGTAAGAATATATATAATATATATATATATATATATATATATATTAAGCTATGGAAC	5248
QY	2247	aaccttlaagatatgtatatttcttgtaattatgatgataaaatttgcattcatgtica	2306
Db	5247	TACAAAAATCAGGATATTATTTCCTTATTTTCATTTTTTAAATAAAATATTATTATTATTA	5188
QY	2307	caacattactttaaaaaatcgtatatttcoactacataatgtgactaagaacaaaaata	2366
Db	5187	TATAGAACGCGATTACATATAAATAAATTAAGTCCATAAATATTCATCAAAAAAACAACAAAT	5128
QY	2367	cattgattattgatataatcgtcaactgaaatttcttcogaggagataaaattctcaaacat	2426
Db	5127	CGAAATTTTCATATGTGAAATGTAGTTAAATTAATTAATAAATGGGACAAATTTTTTTAAT	5068
QY	2427	agcaagaactcctaataatgtttcgtgactacccttttagacgaatttttttaagattcg	2486
Db	5067	GTATTATAAGGAACCTTTTAAGATATAAATAAACATTTTATATAATATATTATTATCCTAT	5008
QY	2487	taacgtgacttatggctc 2504	
Db	5007	TACAGTAACATTTTTTAT 4990	

RESULT 15  
AAAT70223  
ID AAA70223 standard; DNA: 3549 BP.

RESULT 15

AAA70223

ID AAA70223 standard; DNA; 3549 BP.

XX	AAA70223;
AC	
XX	
DT	07-NOV-2000 (first entry)
XX	
DE	Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:356.
XX	
KW	Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine; antimalarial; malaria; protozoacide; infection; insecticide; ds.
XX	
OS	Plasmodium falciparum.
XX	
PN	WO200025728-A2.
XX	
PD	11-MAY-2000.
XX	
PP	05-NOV-1999; 99WO-US26796.
XX	
PR	05-NOV-1998; 98US-0107131.
XX	
PA	(HOFF/) HOFFMAN S.
PA	(CARU/) CARUCCI D.
PA	(GARD/) GARDNER M.
PA	(VENT/) VENTER J C.
XX	
PI	Hoffman S, Carucci D, Gardner M, Venter JC;
DR	WPI; 2000-365347/31.
XX	
PT	Proteins encoded by chromosome 2 of the human malarial parasite,
PT	Plasmodium falciparum, useful as antimalarial vaccines and in the diagnosis of P.falciparum infection -
XX	
PS	Disclosure; Page 545; 577pp; English.
XX	
CC	The present invention describes proteins and their fragments (I) encoded by chromosome 2 of the human malarial parasite, Plasmodium falciparum. Also described are: (I) nucleotide sequences (II) encoding (I); and (2) vaccines against P. falciparum infection comprising (I) or (II). (I) and (II) are useful for the development of vaccines against P. falciparum infection. (I) and polyclonal antisera or a monoclonal antibody raised to immunogens comprising the sequences of (I), are useful in the detection of infection with P. falciparum. Furthermore, (I) (especially when they are rifins or secreted or membrane proteins) can aid the identification of drugs to treat or prevent P. falciparum infection, or they can be used to identify drug resistance in P. falciparum. Sequencing of the Plasmodium chromosome 2 and the subsequent identification of proteins encoded by it will help to expand our understanding of parasite biology, a process hampered by the complexity of the parasitic lifecycle, and provide new targets for vaccine and drug development. Parasite resistance to drugs and mosquito resistance to insecticides have led to a resurgence of malaria in many parts of the world, and there is a pressing need for vaccines and new drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352 represent nucleotide and protein sequences given in the present invention, but which are not specifically mentioned within the specification.
XX	
SQ	Sequence 3549 BP; 1709 A; 352 C; 279 G; 1209 T; 0 other;
	Query Match 2.8%; Score 96.4; DB 21; Length 3549;
	Best Local Similarity 45.6%; Pred. No. 2.8e-06;
	Matches 546; Conservative 0; Mismatches 631; Indels 21; Gaps 5;
Qy	512 aaaatcagaaacagaccagcaatttatgtcttctttataaaaaagaagaagaat 571 
Db	1102 aaaaaaaaaataataataataaaaataataataataataataataataata 1161 
Qy	572 tcattacttatgaatttaaagctaagcagtggaaacctgaaagcagaatttcataa 631 
Db	1162 ataatggaatttattaataataataaggaataataatttcatttttttttta 1221 
Oy	632 aaaaatgaaactgcacaagaacttatgtgtatatataacatcatcataaacaactc 691 

Search completed: July 13, 2001, 00:24:44  
Job time: 7407 sec

```
Db 1222 aaaaaaataataaaaaatgacgtaccataagtttaaaagagaaagatatgaac 1281
Qy 692 aaatatgatataaatacatttttaacaaaaatttctatgaaattataataaaaaaagtcactt 751
Db 1282 acattataatgtgtgataaataataataaataccatcttatttttaataatttt 1341
Qy 752 ttgacacttaaaaggttgacaaataacogtctctcaaaaaaaatcaaaacatttataat 811
Db 1342 caagactctctattttataaaagtataataaaattatttaaaaggcaacataataaat 1401
Qy 812 ttctaaactatggtgtaatttttctgaaatcaaaagaaagaaaggttctctatatcat 871
Db 1402 tatttatgatgatcatggtttttttaagtcatttaagaaa-----tgtgttaaaaa 1454
Qy 872 aagtttcattatgtatcaaaactttcaaaatttcatgttaatttgaaaggaaaaaattaaag 931
Db 1455 gaattgtgcttatttcacagacagggatttgatttatataataatggaaacacacatat 1514
Qy 932 atataatgtgtttttgtttcttatgttacattttctatgggaatataatttcatacaaaa 991
Db 1515 gaataatttgataataataatacaataataataataataaaaaataaataaatacaaa 1574
Qy 992 aatgtattttaatatgtatgagagattaccatccaaaaggttcgaactttataaaaaacaagt 1051
Db 1575 tatgtataataaacagacaaggttaaaagataaacaatgctctatttccctctctc---t 1631
Qy 1052 taataactaaacaatacatatgatacacaatacaatgaacagttttgattcttaaaatagaat 1111
Db 1632 tataaaagatgatatatttcacacattgaagattccattttctcatcataaaagatat 1691
Qy 1112 gatgagcaaacctcaaaaatgtctctttaggatacacaaaatcttctttagcttatta 1171
Db 1692 tatatatatgtttacaacaaaacaaatgtatgaataataaattattccataaaattattaa 1751
Qy 1172 aagccggaggttcaactctctc----tccctgtagactttttgttttcaaatcttttc 1227
Db 1752 ccatttaataacaataataaacaataatgttctaaattatttagtcacaataattatact 1811
Qy 1228 ttccaaaaatcaataatagtttaagggcataatatttggttttaatttaagtccataga 1287
Db 1812 ttgtatacaaaactaaattgtaaaactcagttaaagaaactgttattcactctattataa 1871
Qy 1288 ttttttagaccactctctcaacgacaaaatactctcaaatgttaacacattttaaaactta 1347
Db 1872 taattatagaccactcttaaaacaaagaaataaaagaataataataataataataat 1931
Qy 1348 aagttattgcatcacaactccttaaaatataataataataataataataataataata 1407
Db 1932 ctatttaaaaaataaataaaaaataattataaaaaaaataaaaaaaataataata 1991
Qy 1408 atatatgaagtataatagaacgataactccttactcaacatttagcccaaaaaaac 1467
Db 1992 atatatattatataatgcaaaaaaa-----taagtgtgtaataatacaacaacataa 2048
Qy 1468 atccataatgcatttaactagggaattttaacaaactcaaatagggttggtgtaaaaaa 1527
Db 2049 tgttatgatgacatcaatacaacaatactctatttttagatcccttgagtgtaaaagt 2108
Qy 1528 aaacaaatagtagatgtacatcagtaactttaaataataataactcatatcgaaagtttta 1587
Db 2109 acacaaattattattttaataataattttaataaaaaagtaataataataataata 2168
Qy 1588 aattttgcgaataataataatttatcatcaattaaataacatttaataatgcataatt 1647
Db 2169 atggtccctttatttttttcttcttgattaaacagaaacatgcttttataaaaaaagaag 2228
Qy 1648 ctgtaatatctatcttaatttctcatatagaccacaaataaataaatacaatacaata 1705
Db 2229 gtctatatattatgttatattttt---catatacaaaaataatcataataata 2282
```



**THIS PAGE BLANK (USPTO)**

THIS PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2001, 21:53:12 ; Search time 4439.56 Seconds  
(without alignments)  
11929.464 Million cell updates/sec

Title: PCT-US01-13059-3  
Perfect score: 3424  
Sequence: 1 aagcttaagctaccacat.....aacggcatagctactcagt 3424

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba1: \*  
2: gb\_ba2: \*  
3: gb\_ba3: \*  
4: gb\_in1: \*  
5: gb\_in2: \*  
6: gb\_in3: \*  
7: gb\_om: \*  
8: gb\_ov: \*  
9: gb\_pat1: \*  
10: gb\_pat2: \*  
11: gb\_ph: \*  
12: gb\_pl1: \*  
13: gb\_pl2: \*  
14: gb\_pl3: \*  
15: gb\_pl4: \*  
16: em\_bal: \*  
17: em\_ba2: \*  
18: em\_fun: \*  
19: em\_htgo\_hum: \*  
20: em\_htgo\_inv: \*  
21: em\_htgo\_rod: \*  
22: em\_htg\_hum1: \*  
23: em\_htg\_hum2: \*  
24: em\_htg\_hum3: \*  
25: em\_htg\_hum4: \*  
26: em\_htg\_hum5: \*  
27: em\_htg\_hum6: \*  
28: em\_htg\_hum7: \*  
29: em\_htg\_hum8: \*  
30: em\_htg\_inv1: \*  
31: em\_htg\_inv2: \*  
32: em\_htg\_other: \*  
33: em\_htg\_rod: \*  
34: em\_hum1: \*  
35: em\_hum2: \*  
36: em\_hum3: \*  
37: em\_hum4: \*  
38: em\_hum5: \*  
39: em\_hum6: \*  
40: em\_hum7: \*  
41: em\_in: \*  
42: em\_om: \*  
43: em\_or: \*

44: em\_ov: \*  
45: em\_pat: \*  
46: em\_ph: \*  
47: em\_pl: \*  
48: em\_ro: \*  
49: em\_sts: \*  
50: em\_sy: \*  
51: em\_un: \*  
52: em\_vi: \*  
53: gb\_sts1: \*  
54: gb\_sts2: \*  
55: gb\_sts3: \*  
56: gb\_sy: \*  
57: gb\_un: \*  
58: gb\_vil: \*  
59: gb\_vil2: \*  
60: gb\_htg1: \*  
61: gb\_htg2: \*  
62: gb\_htg3: \*  
63: gb\_htg4: \*  
64: gb\_htg5: \*  
65: gb\_htg6: \*  
66: gb\_htg7: \*  
67: gb\_htg8: \*  
68: gb\_htg9: \*  
69: gb\_htg10: \*  
70: gb\_htg11: \*  
71: gb\_htg12: \*  
72: gb\_htg13: \*  
73: gb\_htg14: \*  
74: gb\_htg15: \*  
75: gb\_htg16: \*  
76: gb\_htg17: \*  
77: gb\_htg18: \*  
78: gb\_htg19: \*  
79: gb\_htg20: \*  
80: gb\_htg21: \*  
81: gb\_htg22: \*  
82: gb\_htg23: \*  
83: gb\_htg24: \*  
84: gb\_htg25: \*  
85: gb\_pr1: \*  
86: gb\_pr2: \*  
87: gb\_pr3: \*  
88: gb\_pr4: \*  
89: gb\_pr5: \*  
90: gb\_pr6: \*  
91: gb\_pr7: \*  
92: gb\_pr8: \*  
93: gb\_pr9: \*  
94: gb\_rol: \*  
95: gb\_rol2: \*  
96: gb\_in4: \*  
97: gb\_pr10: \*  
98: em\_ba3: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3424	100.0	93045	14	ATT32M21
2	191	5.6	2382	13	AF367313
3	180.4	5.3	4601	6	DMU11584
4	180.4	5.3	19517	6	DMU37541
5	174.4	5.1	104992	60	AC005504
6	174.4	5.1	169546	60	AC004157
7	164.4	4.8	67970	96	PFMAL1P3
8	161.8	4.7	14867	4	AE001398

AL162875 Arabidops  
AF367313 Arabidops  
U11584 Drosophila  
U37541 Drosophila  
AC005504 Plasmodiu  
AC004157 Plasmodiu  
AL031746 Plasmodiu  
AE001398 Plasmodiu

FEATURES		Location/Qualifiers	
source	1..93045	misc_feature	676..3112
	/organism="Arabidopsis thaliana"	gene	/gene="T32M21_10"
	/variety="Columbia"	CDS	join(676..841,958..1235,1354..1545,1884..2447,2525..2642,2727..3112)
	/db_xref="taxon:3702"		/gene="T32M21_10"
	/chromosome="5"		/notes="strong similarity to NAC2, Arabidopsis thaliana, EMBL:AF201456"
	1..1000		/codon_start=1
	/note="this clone overlaps with BAC T19M18"		/product="putative protein"
	676..3112		/protein_id="CA85547.1"
	/gene="T32M21_10"		/db_xref="GI:7406445"
	join(676..841,958..1235,1354..1545,1884..2447,2525..2642,2727..3112)		/translation="MGRSVTSIAPGRFRHPTDEELVRYLKRKVCNKPKFQDAISVT DIYKSEPMWDLDPKSLKSRLEWFFSMIDKYSNGSKTNRAATEKYWKTTGKDRIR NGSRVGMKTLVYHKGRAPRGRTNWMHYRLSDLDLKKAGVPOEAYVLCRIFOKS GTGPKNGEOXGAPYLEEWEEDGMTVYPAQDAFEGIALNDVYVDIDDERPNLV YDAVILPNYCHGESSNNVESGNYIQPGNNVYDSGCGYFQPIETFEEDRRKP IIRREGIQPCSLPFEQIGGVQDENNVNLENNNNVFNADTCISDIPIDHNYLPDEP FMDPNNNLPNDGLYLETNDLSCAQDDNFEDYLSFFDDEGLTFDDSLMGPEFLP NQEALDQKPAKRELEKEVAGKEAVEGSGSGSKQDFTDFKDSAPKYPFLKKTIS HMLGAIPTSPSPASQFQTKDAMRLHAAQSGSVHTAGMMRISNMTLAADSGMGWYD KGNLNVLSFGVQVQDDAMTASGSKTGITATRAMLVFMCLWVLLLSVSPKIVTMYSA R"
exon	676..841		/gene="T32M21_10"
	/number=1	intron	842..957
	/gene="T32M21_10"		/number=1
exon	958..1235		/gene="T32M21_10"
	/number=2	intron	1236..1353
	/gene="T32M21_10"		/number=2
exon	1354..1545		/gene="T32M21_10"
	/number=3	intron	1546..1883
	/gene="T32M21_10"		/number=3
exon	1884..2447		/gene="T32M21_10"
	/number=4	intron	2448..2524
	/gene="T32M21_10"		/number=4
exon	2525..2642		/gene="T32M21_10"
	/number=5	intron	2643..2726
	/gene="T32M21_10"		/number=5
exon	2727..3112		/gene="T32M21_10"
	/number=6	exon	complement(3563..3721)
	/gene="T32M21_10"		/number=1
exon	complement(3563..3721)	gene	complement(join(3563..3721,3811..3870,3949..4269,4349..4432,4513..4601,4709..4823,4900..5026,5142..5281,5368..5429,5507..5573,5652..5783,5870..5947,6041..6151))
	/number=1		/gene="T32M21_20"
exon	3563..6151	gene	complement(join(3563..3721,3811..3870,3949..4269,4349..4432,4513..4601,4709..4823,4900..5026,5142..5281,5368..5429,5507..5573,5652..5783,5870..5947,6041..6151))
	/gene="T32M21_20"		/number=1
CDS	complement(join(3563..3721,3811..3870,3949..4269,4349..4432,4513..4601,4709..4823,4900..5026,5142..5281,5368..5429,5507..5573,5652..5783,5870..5947,6041..6151))		

4349. .4432,4513. .4501,4709. .4823,4900. .5026,5142. .5281,  
5368. .5429,5507. .5573,5652. .5783,5870. .5947,6041. .6151))  
/gene="T32M21\_20"  
discoideum, PIR:S68824"  
/codon\_start=1  
/product="putative protein"  
/protein\_id="CAB85548.1"  
/db\_xref="GI:7406446"  
/translation="MEEIRKEIDVGDHNSLAHDEWTPLPVSGSRASARYKHAADVVD  
EKLXIVGSGRGVLDVQVFDLRLSLTWSSLKLTKESSADNIOEDDGSRLREAFPAI  
SDHRMKGWKLILIGHSKSSDNMLVRFIDLETHSCGVTDVFGNVPASGHSITL  
VGSRLVPGGEDKRRLLNDLHLHLERFDVYVETKOTRPVPRFDHTAATHSRYLL  
TFGCGSHSIFYSDLHILDQTMWSQPHVQGVDPFRAGHAGITIDENWYIVGGDINS  
TGCLTELNNKSLVMSSTHVEARHPLASEGLSVCSASVEGENILAFPGYNGKYN  
DIFVMRLKPGSSHPKIFKSPAAAAAASVTAAYAIKSDKSDYPPANPPLNGNS  
LPERDIRNRDITIKEEKRALESSIAETOVENAKLREKIDEVNSSHTLSQLOSVQGO  
LISPSRCFKLEAQIAELQKALSGQSISTEAEEVEMLRQRSASDEEDGTVRQGSAGV  
WGLFGR"  
complement(3722. .3810)  
/number=1  
complement(3811. .3870)  
/gene="T32M21\_20"  
/number=2  
complement(3871. .3948)  
/number=2  
complement(3949. .4269)  
/gene="T32M21\_20"  
/number=3  
complement(4270. .4348)  
/number=3  
complement(4349. .4432)  
/gene="T32M21\_20"  
/number=4  
complement(4433. .4512)  
/number=4  
complement(4513. .4601)  
/gene="T32M21\_20"  
/number=5  
complement(4602. .4708)  
/number=5  
complement(4709. .4823)  
/gene="T32M21\_20"  
/number=6  
complement(4824. .4899)  
/number=6  
complement(4900. .5026)  
/gene="T32M21\_20"  
/number=7  
complement(5027. .5141)  
/number=7  
complement(5142. .5281)  
/gene="T32M21\_20"  
/number=8  
complement(5282. .5367)  
/number=8  
complement(5368. .5429)  
/gene="T32M21\_20"  
/number=9  
complement(5430. .5506)  
/number=9  
complement(5507. .5573)  
/gene="T32M21\_20"  
/number=10  
complement(5574. .5651)  
/number=10  
complement(5652. .5783)  
/gene="T32M21\_20"  
/number=11  
complement(5784. .5869)  
/number=11  
complement(5870. .5947)  
/gene="T32M21\_20"

/number=12  
complement(5948. .6040)  
/number=12  
complement(6041. .6151)  
/gene="T32M21\_20"  
/number=13  
complement(join(7298. .7394,7484. .7545,7615. .7710,  
8136. .8197,8761. .8954,9380. .9486,9592. .9830,10135. .10219))  
/gene="T32M21\_30"  
complement(7298. .7394)  
/gene="T32M21\_30"  
/number=1  
7298. .10219  
/gene="T32M21\_30"  
complement(join(7298. .7394,7484. .7545,7615. .7710,  
8136. .8197,8761. .8954,9380. .9486,9592. .9830,10135. .10219))  
/gene="T32M21\_30"  
/note="similarity to astrocytic NOVA-like RNA-binding  
protein, Homo sapiens, U70477"  
/codon\_start=1  
/product="putative RNA-binding protein"  
/protein\_id="CAB85549.1"  
/db\_xref="GI:7406447"  
/translation="WESPEVYAGSGPELAKRSPHPDSSADSAEKPTHRFLVNSA  
AAGSVIGKGGSTITEFQAKSGARIQLSRNQEFFPCTTDRIIMISSIKEVNGLLEIL  
Query Match 100.0%; Score 3424; DB 14; Length 93045;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 aagcttaagctaccacaatcgaaatttagtaaaagacccatgatttgaaattgggaattgt 60  
|||||  
Db 62465 AAGCTTAAAGCTACCAACATCGAATTTAGTAAAGACCCATGATTTGAAATTTGGAATTTGT 62524  
QY 61 cggcaaaatcgagaagatagcgcacacggaacagtgaaacacacacacacacacgctgtaa 120  
|||||  
Db 62525 CGGCAAAATCGAGAGATATAGAGCCGACACGGGACACAGTGAACACCAAGCGGTAA 62584  
QY 121 gaatgaacagtgaggaggaagagaatcttaccgatcatcttcgagtgaggaagaatgag 180  
|||||  
Db 62585 GAATGAACAGTGGGGAAGAGAGAGAAATCTTACCGATCATTCGAGGAAAAGATGGG 62644  
QY 181 aatcagagaaaaatctggaaaaaagaaattaaagagaaagagagagagaaagtggagag 240  
|||||  
Db 62645 AATCAGAGAAAAATCTGGAAAAAAGAAATTAAGAGAAAAGAGAGAGAAAAGTGAAGAG 62704  
QY 241 gaagatgcagtgaaagactgctatagccacatcccatggtgtgtgatgagagagagagaga 300  
|||||  
Db 62705 GAAGATGCAGTGAAGACTGCTATAGCCACATCCCACATGCTGTGTGATGAGAGAGAGAGA 62764  
QY 301 gagaggttaagcagcaaaattgtggagagataaagagagagagagagagactgagcgagtaag 360  
|||||  
Db 62765 GAGAGTTAAAGCAGCAAAATTTGTGAGAGATTAAGAGAGAGAGAGAGACTGAGCGAGTCAAG 62824  
QY 361 ttcgctgctgtgtttaaagaaagaatccctatatattgctcttttttactactattattt 420  
|||||  
Db 62825 TTCTGCTGCTGCTTTTAAAGAAAGAAATCTATATTGCTTTTCTTTTACTACTTATTT 62884  
QY 421 tcagactatttctctattttgcctcaaaatttttgattgtcacttttcgactcctcaagt 480  
|||||  
Db 62885 TCAGACTATTTGCTTATTTTGGCTCAAACTTTTTTGATTGTGCACCTTTTGCATCTCTAAAGT 62944  
QY 481 gttgacaattttacctgctctttttctccaagaaaaatcagaacacagacacagcaaat 540  
|||||  
Db 62945 GTTTCACAAATTTACCTGCTTTTCTCCAAAGAAAATCAGAACACAGACACACCAATTTA 63004  
QY 541 tgtattttctataaaaaagaagaagaatcattactattatagaatttaaaagctaaag 600  
|||||  
Db 63005 TGTATTTTCTATTAAAAAAGAAAGAAATTCATATTACTATTATAGAAATTAAGAGCTAAG 63064  
QY 601 cagttgaaaacgtgaagcagaaatttctaaaaaaataagtaaaactgctacaaactattt 660  
|||||  
Db 63065 CAGTTGAAAACGTGAAGCAGAAATTTCTAAAAAAATAGTAAACGTCTACAAACTTATTT 63124

Qy 661 atgtgtatataacatatctataaagaaactcaaatatataatgataatcatcttttaacaaat 720  
Db 63125 ATGTGTATATAACATATCTATATAAGAAACTCAATATATATGATGATAATCATTTTAAACAAAT 63184  
Qy 721 ttctatgaaatataataaaaaagtcacttttgacacttaaaaggttgacaataacogt 780  
Db 63185 TTCTATGAAATTAATAAATAAAGTCACTTTTGACACTTAAAGGTTGACAATAACCGT 63244  
Qy 781 ctctccaaaaaaatcaaacattataatttcaaacatgctgaatatttgctgaa 840  
Db 63245 CTCGCCAAAAAATCAAAACATTTATAATTTCTAAACATATGTTGTAATTTTCTGCTGAA 63304  
Qy 841 atcaaaagaaagaggttctcatatcataagtttccatttctgtatcaaaccttcaaa 900  
Db 63305 ATCAAAAGAAAGAGGATTTCTATATCATTAAGTTTCATTATTGTATCAAACTTCAAA 63364  
Qy 901 ttctatgtaattgaagggaaaaaataagataataatgtgtttttgtttctctatgtta 960  
Db 63365 TTTTCATGTAATTTGAAAGGAAAAAATTAAGATATAATGTGTTGTTTGTCTTCTTAAGTTA 63424  
Qy 961 catttcatgaatatattcatacaaaaaaatgtattttaatatataatgagagattacc 1020  
Db 63425 CATTTTCATGGAATATATTCTAACAACAAAAATGTATTTTAATATGATGAGAGATTACC 63484  
Qy 1021 atccaaaaggtcgaaacttatataaaacaagttaataactaaacatacatgctgaccaa 1080  
Db 63485 ATCCAAAAGGTGCAACTTATATAAAACAAGTTAATACTAAACAATACATGTGATCAAA 63544  
Qy 1081 tcaatgacagtttggatcttaaatagaaatgattgagcaaacctcaaaaatgtcttctt 1140  
Db 63545 TCAATGACAGTTTTCATCTTAAAAATAGAAATGATGTGACAAACCTCAAAAATGTCTTCT 63604  
Qy 1141 aggatcacaaactcttctttagcttattaaagccggaggttcaactctctcccttg 1200  
Db 63605 AGGATCACAAAATCTTCTTACCTTATTAAGCCGGAGTTCACTCTCTCTCCCTTG 63664  
Qy 1201 tagacttttcttcaaatctttcttcaaaaaatcaataattagtttaattgggacata 1260  
Db 63665 TAGACTTTTCTTCAAACTTTTCTTCAAAAATCAATAATAGTTAATGGGCATA 63724  
Qy 1261 atatttggtttaataagtcocatagatttttttagggaccatctctaatcagacaaatat 1320  
Db 63725 ATATTGGTTTAAATAAGTCCATAGATTTTATAGGCCACTCTCTAATCAAGCAAAATAT 63784  
Qy 1321 cctaaattgtaacacatttaaaacttaaaagattgtcattcacaaactcttaaaaatatata 1380  
Db 63785 CCTAAATGTAAACATTTAAACTTAAAGTATTGCAATTCACAACTCTTAAATATATA 63844  
Qy 1381 tatatatatatatatatatatatatatatatgaaagtatatagaaacgataaactc 1440  
Db 63845 TATATATATATATATATATATATATATATGAAAGTTATAGAAACGATAAATC 63904  
Qy 1441 cttaactcaaatagccccaaaaaacatccaatgctatttaacttaggaattttaaca 1500  
Db 63905 CTTAATCAAAATTAGCCCCAAAAAATCCATAATGCATTTAAACTTAGGAATTTTAAACA 63964  
Qy 1501 aactcaaataggttggtagtttaaaaaaaacaaatagtagatgtacatacgtacacttaa 1560  
Db 63965 AACTCAAAATAGTTGGTAGTTTAAAAAACAATACTAGATGTACATACGTACCTTTAA 64024  
Qy 1561 aaatatactcatatcgaaagtttttaaattttgcgaaattaaatacatatttatctcaa 1620  
Db 64025 AAATATATACTCATATCGAAAGTTTAAATTTTGCAGAAATTAATAACATATTATCTATCAA 64084  
Qy 1621 ttaaaatcattttaaatgataatctgataatctatctttaaatttccatataaac 1680  
Db 64085 TTAATAATACATTTAATGATGAATTTCTGTAATATCTATCTTTAATTTCCATATAGAAC 64144  
Qy 1681 caaaacaaaataaacatatcaaatagtttttaacttaacaaaaacggttaggaaaaagtta 1740  
Db 64145 CAAAACAAAATAAACATATCAATAGTTTAACTTAACTAAACAAACGTTAGGAAAAAGTTGA 64204

Qy 1741 cctaactagctgttgaacttgtaactgtcaatcgaaagcgatatttccaatatatac 1800  
Db 64205 CCTAACTAGCTTGTATGACGTGAACTTGCAATGCGAAAGCGATATTTCGAATATATAC 64264  
Qy 1801 tacaatgtagtattatttatattggaagtttctaaaaaggtgtgtgagttggttacttgt 1860  
Db 64265 TACATGTAGTATTAATTTATATGGAAGTTTCTAAAAAGGTGTTGAGTGGATTGTACTTGT 64324  
Qy 1861 tggaggtgctattttctctcttgcataataattttacgagtatgggataaactacata 1920  
Db 64325 TGGAGGATGCTATTTTTTCTCTTCCCTTCTGCAATATATTTTACGAGTATGGGATAACTACATA 64384  
Qy 1921 ctcatgattatgaacgcctcaactttattgaaaaacctctcaatacaccaaaatgtcac 1980  
Db 64385 CTCATGATTATGAACGCTCACTTTATTGAAAAACCTCTCTAATACACCAAAATATGTACAC 64444  
Qy 1981 tagattccaaaacgtagaccaaattgtatctaattcctaattctcaatacaaaagttaatt 2040  
Db 64445 TAGATTCCAAAACGTAGACCAAAATTTGTATCTAATCTCAAAATCTCAATCAAAAGTATTAAT 64504  
Qy 2041 taccgatggttaagaaaaagtttaaccgatataattatcaaaaagaaagataaagtcacacagat 2100  
Db 64505 TACCGATGGTAAGAAAAAGTTTAACCGATATATATTAACAAGAAAGATATAGTCAACAGAT 64564  
Qy 2101 tcttaactctcttatttttgggtatagaaacatttgcacaaaactctcaaaagatatgaac 2160  
Db 64565 TCTTAATCTCTTTTATTTTGGTATATGAACATTTGTACAAAAATCTCAAAAGATATGTAAC 64624  
Qy 2161 tgtttaaaaataaaaattcaactgagatttaattcttcagactcgtgttagctataataatgt 2220  
Db 64625 TGTTTAAAAATATAAATTTCACTGAGATTAATTTCTTCAGACTCGTGTAGCTATAATAATGT 64684  
Qy 2221 caagagttctcttcttcttcaagaaaaaaccttaagatatgtatatcttctgtaattatg 2280  
Db 64685 CAAGAGTCTCTCTCTGTTTCAAGGAAAAACCTTAAAGATATGTATATTTTCTGTATTAATG 64744  
Qy 2281 atgataaatttgcatttcttgcacaaacattactttaaanaactcgtattttcattac 2340  
Db 64745 ATGATATAAATTTGCTATTATTCATTGTCACAAACATTTACTTTAAAAAATCGTATTTTCATTAC 64804  
Qy 2341 tacaattgtgactaagaacaaaaaacatcatgatttattgatatatcgtcaactgaatttc 2400  
Db 64805 TACAATGTTGACTAAGAACAAAAATACATTGATTTATGATATATCGTCAACTGAATTTTC 64864  
Qy 2401 ttccgagggataaattctcaaacatagcaagaatctcataataatgtttctgtaactacc 2460  
Db 64865 TTCCGAGGGATATAATTTCTCAACATAGCAAGAAATCTCATATAATGTTTCGTGACTACC 64924  
Qy 2461 tttagacgaaatttttttaagattcgttaacgtgacttatggtctctctgctgtgggggtca 2520  
Db 64925 TTTAGACGAAATTTTAAAGATTCGTAAGCTGACTTATGCTCTCTTGTCTGTGGGGTCA 64984  
Qy 2521 atcgaaataatctaaatgtatggagtcataataaaataccaagaaaaataaaggagcag 2580  
Db 64985 ATGCGAATAAATCTAAATGATGGAGTCAAAATAAATAACCAAGAAAAATAAAGGAGCAG 65044  
Qy 2581 caccataaactcataaggaccagaaactcttcttcatgtgtttaaaaataggattatcccc 2640  
Db 65045 CACCCAATAAATATATGCGACCCAGAAATCTTTTCATTTGTTTAAAAATAGGATTTATCCCG 65104  
Qy 2641 aagagtagaaggactaaattgaaactgattggggtaggaagagacatccgtcacaaatcata 2700  
Db 65105 AAAGATGAAGGACTAAATTTGAACTGATTTGGGGTAGGAAGAGATCCGTGCACAAATCATTA 65164  
Qy 2701 atggctccacgcggaactgtcgtttatcaatttcatcttaacttcggtcggtgttta 2760  
Db 65165 ATGGCTCCACGCGGAACCTTGCGTTATACAAATTTCAATTTACTTTTCGGGTTCGGGTTTA 65224  
Qy 2761 tattccaaatgggtcacaataatattagttttaatacacactaacgagtaataatttgggtgac 2820  
Db 65225 TATTCCAAATGGGTCAAAATTAATTAGTTTAAATACACTAACCGAGTAAATTAATTTGGTGAC 65284  
Qy 2821 tacaatttttatacgtttgtgcaatttagaacaacagacaatagtcgtaaaaaacacgagttcgg 2880













Db 18075 ACACAAAAATTTTAAATAAATAAATTTTAAATGAATAATAATTTTATTTATTTTCAATTT 18134  
 Qy 1946 attgaaacacctcctaataacacacaaatgtcactagattcctcaaaacgtagacaaattg 2005  
 Db 18135 TTTTAAAAATTTTAAAAAATAAATTTTAAAAAATAAATTTTAAAAAATAAATTTTAAAA 18194  
 Qy 2006 tatctaactcacaattcctaacaagattaaatttaccgtaggttaa-----gaaagt 2059  
 Db 18195 TAAATTAATAGATATTTAT 18254  
 Qy 2060 taaccgatataattatcaaaagaagaalaagtcacacagattccttaactctcttatttg 2119  
 Db 18255 TTAATAAATAAATAAATTTAAAAATTTAAAAATTTAGATATATATTTATATAAATAATTTATATCTC 18314  
 Qy 2120 gtatagaacatttgcacaaaatcacaagaatgtaactgttttaaaataataaattca 2179  
 Db 18315 ATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 18374  
 Qy 2180 ctgagattaaattcttcagactcgtgttagctataataatgtaagagttcttctgttttc 2239  
 Db 18375 ATATTTATTAATTTATATATATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 18432  
 Qy 2240 agggaaacacccctaagaatgataatttcttctgttaatttatgatataatttgcattc 2299  
 Db 18433 AATTAAATTTTAAATAAATTTAAATAAATTTTAAAAAATAAATTTTAAAAAATAAATTT 18492  
 Qy 2300 attgtcacaaacattcaacttcaaaatcgatttttcttcttcttcttcttcttcttctt 2359  
 Db 18493 TTATAAAAATTTATTTATTAATAAATTTATTTATTTATTTATTTATTTATTTATTTAT 18552  
 Qy 2360 aaaaatacattgatttgatataatcgtcaactgaatttcttccgagggtataaattct 2419  
 Db 18553 ATAAATAAATTTTAAATGAATAATAAATTTATTTATTTATTTATTTATTTATTTATTT 18612  
 Qy 2420 caaacatgacgaagaatcctcat---aataatgttctgtgactcactccttgcagaaatttt 2476  
 Db 18613 TTAATAAATAAATAAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 18672  
 Qy 2477 ttaagattcgtaactgacttctgctctgctgtgctgtgctgtgctgtgctgtgctgtg 2531  
 Db 18673 TTTATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 18732  
 Qy 2532 -tctaattgtatggagtcacataaataaccacgaagaaataacagagagcagcaccacataa 2590  
 Db 18733 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 18792  
 Qy 2591 actatattggaccagaatccttcttcttcttcttcttcttcttcttcttcttcttctt 2650  
 Db 18793 TTAATTTAATTTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 18852  
 Qy 2651 gactaaattgaaactgattggggtaggaagagatccgctcaacatcattaatggcttcca 2710  
 Db 18853 ATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 18912  
 Qy 2711 cgcggaaactgtcgttttatacacaatttcaacttcttcttcttcttcttcttcttctt 2770  
 Db 18913 TAAATTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 18969  
 Qy 2771 gggcacaataatttagtttaatacacactacacgagagtaatttaattggtgactacaatt 2830  
 Db 18970 TTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 19029  
 Qy 2831 tcagtttgggtgcaatttagaacaacacatagtcgtaaaatacaggttcggtgttacatt 2890  
 Db 19030 ATAATGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 19089  
 Qy 2891 tatttaccgttaaaaaatacagaaatttcttcttcttcttcttcttcttcttcttcttct 2950  
 Db 19090 ATTTTATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 19149  
 Qy 2951 atggaataatttagactctagcgaataatgatttcttcttcttcttcttcttcttcttct 3010  
 Db 19150 AAATATTTAATAATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 19209

Qy 3011 tttagcattgaacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3070  
 Db 19210 ATATAG-ATATAAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 19268  
 Qy 3071 gcaacattaaatttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 3130  
 Db 19269 ATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 19327  
 Qy 3131 agaccacaaacattaaatttcttcttcttcttcttcttcttcttcttcttcttcttctt 3190  
 Db 19328 AATATTTAT 19387  
 Qy 3191 ccacaaaagtgggttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3250  
 Db 19398 AAAATTTAAAAAATAAATTTCTTAAATGATATTTTAAATAAATAAATAAATAAATAA 19447  
 Qy 3251 taataggactaatgacataacacaaataacacataacacataacacataacacataacac 3310  
 Db 19448 AATTTAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 19507  
 Qy 3311 taattt 3316  
 Db 19508 TTATAT 19513

RESULT 5  
 AC005504  
 LOCUS Plasmodium falciparum chromosome 12, \*\*\* SEQUENCING IN PROGRESS  
 DEFINITION \*\*\* 3 unordered pieces.  
 ACCESSION AC005504  
 VERSION AC005504.3 GI:4558584  
 KEYWORDS HTG; HTGS-PHASE1.  
 SOURCE malaria parasite P. falciparum.  
 ORGANISM Plasmodium falciparum  
 Eukaryote; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.  
 REFERENCE 1 (bases 1 to 104992)  
 AUTHORS Hyman, R.W., Fung, E.L., Qin, F., Tamaki, T., Kurdi, O.B., Conway, A.B. and Davis, R.W.  
 TITLE Plasmodium falciparum 3D7 chromosome 12  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 104992)  
 AUTHORS Hyman, R.W., Qin, F., Fung, E.L., Conway, A.B. and Davis, R.W.  
 TITLE Direct Submission  
 JOURNAL Submitted (21-AUG-1998) Stanford DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA  
 COMMENT On Apr 2, 1999 this sequence version replaced gi:4337172.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 3 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 58642: contig of 58642 bp in length  
 \* 58643 58842: gap of unknown length  
 \* 58843 91011: contig of 32169 bp in length  
 \* 91011 91212: gap of unknown length  
 \* 91212 104992: contig of 13781 bp in length.  
 FEATURES  
 Location/Qualifiers  
 1..104992  
 /organism="Plasmodium falciparum"  
 /db\_xref="taxon:5833"  
 /chromosome="12"  
 BASE COUNT 44286 a 9326 c 9564 g 4141 t 405 others  
 ORIGIN

Query Match 5.1%; Score 174.4; DB 60; Length 104992;  
 Best Local Similarity 45.8%; Pred. No. 4.2e-14;

[illegible]

LOCUS	AC004157	169546 bp	DNA	HTG	12-AUG-2000
DEFINITION	Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN PROGRESS ***, 2 unordered pieces.				
ACCESSION	AC004157				
VERSION	AC004157.8	GI:9797712			

KEYWORDS

HTG; HTGS\_PHASE1.

SOURCE

malaria parasite P. falciparum.

ORGANISM

Plasmodium falciparum

REFERENCE

Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

AUTHORS

1 (bases 1 to 169546)

TITLE

Hyman,R.W., Fung,E.L., Qin,F., Rowley,D., Mao,J., Tamaki,T.,

JOURNAL

Kurdi,O.B., Conway,A.B. and Davis,R.W.

REFERENCE

Plasmodium falciparum 3D7 chromosome 12

AUTHORS

Unpublished

TITLE

2 (bases 1 to 169546)

JOURNAL

Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W.

AUTHORS

Direct Submission

TITLE

Submitted (19-FEB-1998) Stanford DNA Sequencing and Technology

JOURNAL

Center, Stanford University, 855 California Avenue, Palo Alto, CA

COMMENT

94304, USA

On Aug 12, 2000 this sequence version replaced gi:8810447.

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 2 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1

\* 23466: contig of 23466 bp in length

\* 23467 23666: gap of unknown length

\* 23667 169546: contig of 145880 bp in length.

\* Location/Qualifiers

1..169546

/organism="Plasmodium falciparum"

/db\_xref="taxon:5833"

/chromosome="12"

/clone="PFVAC293"

/clone="3D7"

BASE COUNT 69871 a 15381 c 15705 g 68389 t 200 others

ORIGIN

Query Match 5.1%; Score 174.4; DB 60; Length 169546;

Best Local Similarity 45.8%; Pred. No. 3.9e-14;

Matches 927; Conservative 0; Mismatches 1061; Indels 36; Gaps 8;

QY 414 ttattttcagactattgtcttattttgctcctcaaaacttttttgatgtcacttttcgac 473

DB 84191 TATATTAATAATAATTAAGTTAATTAATTAATAATAATAATAATAATAATAATA 84250

QY 474 ctcaagtggtgacaatttacctgctttttctcagaagaaatcagacagaccagc 533

DB 84251 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 84310

QY 534 aaatttatgtattttcttataaaagaaagaaagaaatcttacttatagaatttaa 593

DB 84311 ATAAATAGAAAAATTAATACAAATTAATTAATTAATTAATTAATTAATTAATTA 84370

QY 594 agctaagcagttgaaacgctgaaagcagaatttctcaaaaaaaatagtaaacctgtacaaa 653

DB 84371 AAATAATAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 84430

QY 654 cttatttggtgataacatactctataagaaactcaaatatataatgataaattcattta 713

DB 84431 TTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 84490

QY 714 acaaaattctatgaattataaaagaaagctcacttttgacacttaaaaggttgacaa 773

DB 84491 ATAAATTTTAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 84550

QY 774 taacgctctccaaaaaaatcaaacatttbaaattttctcaaaact---atgggtgtaa 830

DB 84551 AATAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 84610

QY 831 ttgtgctgaatacaaaagaaagaggttctctatcataagtttcatttgatca 890

DB 84611 ATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 84670

QY 891 aactttcaaatctcatgtaatttgaaagggaaagaaataaataatgattgtttttgtt 950

DB 84671 TATAAATTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 84730

QY 951 tcttatgtttacattttctcatggaataatataatacaaaaaaaatgatttttaataatgatg 1010

DB 84731 AATTATATATTAGTAAATAAAATA-AAATTAATAATTAATAATAATAATAATAATAATA 84789

QY 1011 agagattaccctccaaaggctcgaaacttatataaaacaagtttaatacaacaacatcat 1070

DB 84790 AAAAATAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 84849

QY 1071 gtgatccaaatcaatgacagtttgatctttaaataagaaatgattgagcaaacctccaaa 1130

DB 84850 TT-----AAATATATATACAAATTAATAATAATAATAATAATAATAATAATA 84903

QY 1131 atgtcttttaggacacaaatctttctttagcttatataaagccggagttccaactct 1190

DB 84904 ATAATTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 84963

QY 1191 ctctccctgtgacactttttgttttccaaatcttttcttccaaaaaatcaataatagatt 1250

DB 84964 TAAATATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85023

QY 1251 aatgggcataataatttggttttaagtcacatagatttttttaggaccatctctaatca 1310

DB 85024 ATATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85083

QY 1311 cgcaaatatocctaaattgttaacacatttaaaacttaaaagttgattgcattcacatcctt 1370

DB 85084 TAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85143

QY 1371 aaatagaa 1430

DB 85144 AAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85203

QY 1431 acgataactcttactcaacaaatagcccaaaaaaacacatccatcattgatttaaaactagg 1490

DB 85204 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85262

QY 1491 aatttttaacaaactcaaataggttggtagtttaaaaaaaacaaatagtagtagtagtacatc 1550

DB 85263 TATTATATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85322

QY 1551 gtacctttcaaaatatactacatcagaaagtttttaaaatttgcgaaattcaaacacatt 1610

DB 85323 AATTAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85382

QY 1611 tatctatcaattaaataacatttaataatgcataatctctgtaatatctctatcttcaatttc 1670

DB 85383 TATATATTAATTAAGTTAATTAATAATAATAATAATAATAATAATAATAATAATAATA 85442

QY 1671 catatagaacaaacaaacaaacaaacaaacaaacaaacaaacaaacaaacaaacaaacaa 1730

DB 85443 TATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 85500

QY 1731 gaaaagttgacctgaactgactgactgactgactgactgactgactgactgactgactgact 1790

DB 85501 -----TATTTTAAATTAATAATAATAATAATAATAATAATAATAATAATAATA 85548

QY 1791 caatatatactacatgtagtattatttatatttggaagtttcttaaaaaaggtgttgagtgat 1850

DB 85549 ATAAATAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85608

QY 1851 tgttactgttggaggatgctatttttctctctctgcaataatattttaoagagatgggga 1910

DB 85609 TATATTTAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85668

QY 1911 taactacatcatgattatgaacgcctcacttttatttgaacaaacctcctaata-cacc 1969

DB 85669 TTAATATATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85728

```
QY 1970 aaatgtcactagattccaaacgtagaccattgtatcttaattcctcaaatctcaatca 2029
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85729 TAATATTTATATATTTGTTTAAATATTTAAATTTAAATTTAAATTTAAATTT 85788
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2030 aagtattaattaccgatgtgaagaaagttaacgagataataattcaaaagaagaata 2089
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85789 AATTTATTTAATATATATTTTAAATAAAATATATATATTAATGTAATAAATA 85848
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2090 agtcaacagattctt-----aattcttatttgggtatatgaacatttgcataaaat 2143
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85849 AATTAAGAAATTTATTTAAATAATATATGTTTAAATTTATTAATAATATATATAT 85908
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2144 ctcaaaagatgttaactgttttaataataataattcaactgagattactcttcagactcgt 2203
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85909 TTATTAATATTTTAAATTTAATGTTTATTTATATTAATTAATTTAATTTAAT 85968
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2204 gttagctataaagtcaagagttcttctgttcaaggaagaaaccttaagatatgta 2263
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85969 TATTATTTTAAATATTT-----ATTTAATTTAATTAATTTAATTTAATTTT 86024
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2264 tatttctgtaattgatgatataatttgcattcattgtgcacaaacattactttaaa 2323
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 86025 TATTTTAAATTAATATATATATTAATTTAATTTAATTTAATTTAATTTAATTTA 86084
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2324 aatcgtatttttactactacaatgttgactgaagaacaaatacatattgatttatata 2383
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 86085 TAATAAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTT 86144
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2384 tegtcaactgaatttcttcgaggagataataattctcaaacata 2427
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 86145 TTTTAAATTAATTAATTAATTAATTAATTTAATTTAATTTAATTTAATA 86188
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 7
PFMALIP3
LOCUS PFMALIP3 67970 bp DNA INV 15-DEC-1999
DEFINITION Plasmodium falciparum MALIP3, complete sequence.
ACCESSION AL031746
VERSION AL031746.9 GI:6594243
KEYWORDS HTG.
SOURCE malaria parasite P. falciparum.
ORGANISM Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 67970)
AUTHORS Bowman,S., Church,C., Harris,B., Harris,D., Lawson,D., Quail,M.
and Barrell,B.
TITLE Direct Submission
JOURNAL Submitted (24-SEP-1998) P.falciparum Genome Sequencing Consortium,
The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA, UK
COMMENT On Dec 16, 1999 this sequence version replaced gi:5763807.
For more information about this sequence or the Malaria Project,
see http://www.sanger.ac.uk/Projects/P.falciparum. IMPORTANT: This
sequence is unfinished and does not necessarily represent the
correct sequence. Work on the sequence is in progress and the
release of this data is based on the understanding that the
sequence may change as work continues. The sequence may be
contaminated with foreign sequence from E.coli, yeast, vector,
phage etc.
FEATURES
Location/Qualifiers
1..67970
/organism="Plasmodium falciparum"
/strain="3D7"
/db_xref="taxon:5833"
/chromosome="1"
complement(join(1748..2598,2748..2848,2990..3276))
/gene="MALIP3.01"
complement(join(1748..2598,2748..2848,2990..3276))
/gene="MALIP3.01"
/note="MALIP3.01, conserved hypothetical protein, len: 412
aa, similarity: UPF0006 family eg to
YBL055C/YBL0512/YBL0511, YBF5_YEAST (418 aa), fasta
scores: opt: 316, E(): 1.1e-12, (33.2% identity in 271 aa
```

```
overlap")
/codon_start=1
/product="conserved hypothetical protein, UPF0006 family"
/protein_id="CAB63556.1"
/db_xref="GI:6594244"
/translation="MKLVFVHKYINVLFYISIFLKSLSKLIYNDLRYISTVNNKYK
LOIKRSLKKNHNIRKMDNEDSSFDIGSLNLTOKMPDGVYNSKKHNDLONLNRK
NNNDKIIITCTCLAEIDKSLKICETYPBEGKFLYLSAGVHTNCYEFIDNKHEEKE
IIAKKEYEEFIYFKNEQVENSVMKNGKICDGEKDMNNLNEILLENKDOTIPGFY
NEKDEYLENLKNIKIYPNIRVICIGIEGLDFDRLYFCSYKIQYFQLKQVOMFN
LPMFLHMRNCSETFQIVDIYKFLFKNGGVIHSTFKDEIVHIIVQNYKNLYIGVNG
CSLSLENIYAVKKIPLNLLLETDAPWCGVKKTHASVEYIKDTYKRAYTNLKKIK
LIKDDNTIFKERNPEYNIA"
complement(2599..2610)
/note="potential splice acceptor sequence"
complement(2742..2747)
/note="potential splice donor sequence, atg/gttaa"
complement(2849..2861)
/note="potential splice acceptor sequence"
complement(2984..2989)
/note="potential splice donor sequence, aaa/gtaaaa"
5005..5496
/gene="MALIP3.02"
5005..5496
/gene="MALIP3.02"
/note="MALIP3.02"
/note="MALIP3.02, hypothetical protein, len: 163 aa,
contains possible signal sequence"
/codon_start=1
/product="hypothetical protein, MALIP3.02"
/protein_id="CAB63557.1"
/db_xref="GI:6594245"
/translation="MKLLNRFVLCPIIILFFFLNSVVLGNNNRNNNFHETENAAK
AMRKLSEINSIKLDGDELKIKINDEKHDKSTKWDKSYFISLEEEKYSQDLEFR
KKQINEANTKIIEDROBEFYILNDEIENIATRFVLENNFDELYIQSFQSLIDIOS
LN"
8020..10389
/note="possible cen1, region of very high [A+T] content"
14884..20352
/gene="MALIP3.03"
14884..20352
/gene="MALIP3.03"
/note="MALIP3.03, putative ABC transporter, len: 1822 aa"
/codon_start=1
/product="putative ABC transporter"
/protein_id="CAB63558.1"
/db_xref="GI:6594246"
/translation="MTYKENVGSKGKKKKKQNTSFLNLSFDFWIRPLINDLIK
GDIOPLNCRNFDVPYVASKLEENLRDIEVEDSEYSEKSNSEHVLHGNDSASE
KKNVYVYHNLWSILTKTFRILIIISFYILETLVTLGCKFDIYWRILEGOKIPV
YISLKDPKVFSGLVYVMIMFFHLFEALLHFYFHLFTLNKLSYLMYFLYKINCSNN
NHLQNPDAFYNTYRKFSSTOIDEISRDFLSIGKNASSSSGSKNNNNKNDNNKFEVN
DYIINFIKTKMEKDSINERSLPNVNIYIMFSDPSPVTFVTCINLNFVFKIF
MSYVFHIGKSNVGIWIALSYALISAMILFELPSLFKSLYLRDKRIDNMHVL
KEFKLIKFNWESFAFYINIRFMKMKYCKIRLVLSNIGYFISISDIDIVEVTFEI
YKDLAKKKEIKETSIMPLYVYKILISVANFPLVNMVWGIYVNLKRLNYYINDH
LYNDKLMFYRTYRINEDYINIVDTFLQNETISHDGDSHDKHAKVINKLNT
MKFYFFYKMYNHKNIINQKISGLKANVDNDNKKIKFQEHKSNSTNYNSSHIHE
KKEYENIHNSNSTMSNEFKKKKNNEYIIKLENCISGLSYDNKNDNDHILKNFN
LKRNSLIIIGNVSGSAFHSILGDNFMTHGLYIENFFKMPILVVPONSMLFPM
NIRSMILPGEYNPILYKYTILOQELLANDSTIEHGMKYINDDNHLSKGOKVRICLA
RALYHYIHMHLCTDYEKKLIQNEILDKDLINNNKISSYNNKSKLVNIPNEN
YLQKCLMDNNFYLYLDDIDFTSKISKIFSNLCKEDNISFKDNCSTFISNNKS
TLDNFLIEDILDNFOYEVNIFEIDOKTLKRGNTSEYMEKNLNTKESHWGYSNLNT
IDYTRIKLDFEELNHNKSNMIIYKAYFVKGNTSEYSEFIDSTNDEYIKKMKKNY
KHEHMKNNKDNNNNNKDDHINIMNDNHRNYNDNLGPNSTDSPVSSIGNE
YTLDTYTSNNSDKBEIVKPLYKDTHEENKSSMPFVKSSSNMNPNSFYEDNSSS
FKGSLISLETYVILPISLISGICFSMIYITSAIKVHNHNLISILNAPLIYF
TILQKQRYLEYFVILPISLISGICFSMIYITSAIKVHNHNLISILNAPLIYF
YNNLGNINRFFIIDSADFGLKRYKAFPIFFRCFSLLSLIYIMYRDCITFIPFV
IILYFFYFKFCKEGRKQRLSCHPTPLQNIYSNALSNGKNIINIKYKKNVRI
HYNNFRISYFFKWLINIWASLYIKIFILLTTYIIMHPHLYASGIILKYKEKNVRI
LSTLGYCISFARGLGVIKFLLCDYTHIEKEMCCVQRELEFAKISKNENAKNEL
NVITQTITKERNENISDLSAIVEYKYNVLSIINSSQDDSKKYGKFNENVYSYK
```







```
QY 1146 cacaaaaattcttcttgagcttattaaagccggaggtccaactctctctcccttgtagac 1205
Db 8664 AATTAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATAATAA 8723
QY 1206 ttcttggttccaaactcttcttccaaataataatagtttaagggcacaataatt 1265
Db 8724 TATATTAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATA 8783
QY 1266 tgggtttaattgaagtcctagatgtttttaggaccatctcttaacacgacaaatacctaa 1325
Db 8784 TATATTAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATA 8843
QY 1326 attgtaacacattttaaacttaaaagtagtcattccaaactccttaaaatatatatat 1385
Db 8844 AATTAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATAATA 8903
QY 1386 atatatatatatatatatatatataatgaagttatatagaaacgataactcctcac 1445
Db 8904 AATTAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATAATA 8963
QY 1446 tcaacaattagcccaaaaaacacatccataatgcatttaaaacttaggaatttttaacaaactc 1505
Db 8964 AAAAAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATAATA 9023
QY 1506 aaataggttggttagttaaaaaaaacaaatagtagatgtacacgtagcctcttaaaaaata 1565
Db 9024 TTTAATAATATTCCTCTTATTTATTTAATAATAAATAATAAATAATAAATAATAAATA 9083
QY 1566 tatactcatatcgaaagttttaaatttgcgaaataataatcatatcatcaatcaaa 1625
Db 9084 TAATAATTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 9143
QY 1626 atacatttaataatgcataattctgtaatatctatctctttaaatttccatatagaacccaaa 1685
Db 9144 TATAATAATTTAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATAATA 9203
QY 1686 caaaataaacatatacaaatagttttaaacttaacaaaaacggttagggaaagtgacctaa 1745
Db 9204 AATAATAATTAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATAATA 9263
QY 1746 ctagcttgattgacgttgaactgtccaactgccaagcgagatatttccaatatatactacat 1805
Db 9264 ATAATAATAAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATAATA 9321
QY 1806 gtagtattatttatatgaaagtttccaaaggttgcagtgagtggttacttggtgag 1865
Db 9322 TTAGTAACATTTAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATAATA 9381
QY 1866 gatgctattttctctcttgccataataattttacgagtagtagtggaataactacatactcat 1925
Db 9382 AAAACAATAATATATTTATTTATTTTAAATAATAAATAATAAATAATAAATAATAAATA 9441
QY 1926 gatttgaacgctcacctttatttgaacacccctcctaatacacccaaatat----- 1975
Db 9442 TTTTATTAATAATATTTTAAATAATAAATAATAAATAATAAATAATAAATAATAAATAATA 9501
QY 1976 gtaactagattccaaacgtagaccaatgtctctaactcctaactcctaactcctaactc 2035
Db 9502 ATTATAACCTTTAATAATAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATA 9561
QY 2036 taatttaccogattgaagaaggttaaccgataataattatcaaaagaaagaataagtcac 2095
Db 9562 ATAATAATAATTAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATA 9621
QY 2096 cagattcttaactctcttattttgttgatgaacatttgcacaaatact-----ca 2147
Db 9622 ATATATATATATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 9681
QY 2148 aaagatatgtaactgttttaaaataataaattcactcagatgaatattctcagactcggtta 2207
Db 9682 TATAATATATATATTTTATTAATAAATAATAAATAATAAATAATAAATAATAAATAATAAATA 9741
QY 2208 gctataataatgtcaagagttctctctgttccaaaggaacaccccttaagagatgatatatt 2267
```



Db	7832	TTTAATTAATCTATTTGTTGATGTAATAATAA--AACATATTATATTTAAATATATAAA	7889
QY	1335	catttaaaacttaaaagtattgcattcacatcccttaaaatataatatatatatatata	1394
Db	7890	TGCGTATAATAATAATCTCTTTTAATTTATTTTATAATAATAATTTATTTATTA	7949
QY	1395	tatatatatatatatatagaagttatagaaacgataaactccttactcaacaatt	1454
Db	7950	ATTATTATTATATATATAATTTATTTTAAATGCTTAATAATAATAATAATAAT	8009
QY	1455	agcccaaaaaacatccatcattgaacttaggaatttaa--caaaccaaataggg	1512
Db	8010	-----AATATAAATTAATTAATAATAATAATAATAATAATAATAATAATA	8064
QY	1513	tbggtagttcaaaaaaacaatagtagtgatcacacgttaacottttaaataatactc	1572
Db	8065	TATAAATAATATATATATATATAATTAATTTATTTTATTTTAAATGAATAATAA	8124
QY	1573	atatcgaaagttttaatttgcgaattcaatacacatttctctcatcaataaaacatt	1632
Db	8125	TAAATTAATTAATTAATTTAATTTAATTAATAATTTATTTATATATTTAAATTAAT	8184
QY	1633	taataatgcataattctgtaattctatctctttaattccatagacccaaacaaata	1692
Db	8185	AATATATTTATTTTAATAAACTTAATTAATTAATTTATTTATTAATAATAATA	8244
QY	1693	aacatatcaaatagttttaacttaacaaacgcttagggaaagttgaocctaactagctt	1752
Db	8245	AATTATGATATATATATAATATTAATATATTTTAATTAATTAATAATAATAAT	8304
QY	1753	gattgacgttgaaactgtcaatgcgaacgcatatttccaatatatactacatgtagtat	1812
Db	8305	ATTAAATTAATTTTAAATAAATTAATTAATTAATTAATTAATTAATTTATTTT	8364
QY	1813	tatttatacgaagttctcaaaagcgttgtagtgattgttacttgctggaggatgcta	1872
Db	8365	TTATAATTAATTAATTAATAATTAATTAATTAATTAATTTGTTCTTGTATTAATA	8424
QY	1873	tttttccctctgcataaatatttaccagatgggatacacaatactcatgattatg	1932
Db	8425	TATTTATTTATTTTATTTATTTAATAAATAAATAATTAATAATAATAATAATA	8483
QY	1933	aaacgtcacttatttgaaaacctcccaatacacacaaatagtctacatgattcaaaa	1992
Db	8484	TTATTTCATTTATTTAATAATATATTATTAATAATAATAATTTCCAAATATAATAAT	8543
QY	1993	cgtgaccaaattgtatcctaactcctaattctcaatcaaatgaatttaattaccgatgtaa	2052
Db	8544	AATGTATTTATTAAGCTTAATAATAATTAATTTATTTTATTAATAATAATAATTCGTT	8603
QY	2053	gaaagftaacccgataataattcaaaaagaagaaagaacacagagattcttaactctt	2112
Db	8604	AATTCATTAATTAATAATTTATTTATTTTAAAAATAATTAATGATATAATAATATA	8663
QY	2113	tatttgggtatgaacatttgtacaaaactcctcaaaagatatgttaactgtttaaaatat	2172
Db	8664	TATATTAAATAATTAATTAGTATATAAAAATAATAACATATACATTAATAATTACA--CA	8721
QY	2173	aaattccactgagatgaattctcagactcgtgttagctataataatgtcgaagctcttc	2232
Db	8722	TAAATTTATTTATTTTAAATAATTTTACTTATTTTAAATTAATTAATTTATTTAT	8781
QY	2233	tgtttcaggagaaaaccttaagatatgatatttctgtgaattatgatgataaattt	2292
Db	8782	ATTATTATTTATTTATTTAATTAATTTATTTAATAATAATAATTTATTTATTTATTT	8841
QY	2293	gctattcatgtcacaaaacattactttaaaaaaatcgcattttcattactacaatgct	2349
Db	8842	ATTATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	8898
RESULT 9			
AE001398/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REMARK			
REFERENCE			
AUTHORS			
JOURNAL			
FEATURES			
source			
gene			
CDS			
gene			
CDS			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REMARK			
REFERENCE			
AUTHORS			
JOURNAL			
FEATURES			
Location/Qualifiers			
1. .14867			
/organism="Plasmodium falciparum"			
/db_xref="taxon:5833"			
/chromosome="2"			
complement(1570. .2424)			
/gene="PF0490c"			
complement(1570. .2424)			
/gene="PF0490c"			
/note="Predicted by GlimmerM"			
/codon_start=1			
/product="hypothetical protein"			
/protein_id="AAC71887.1"			
/db_xref="GI:3845198"			
/translation="MKEKEKIMDYLSCPDLDVDVREKSGKNSLLKSSSTKKSDDYKK			
SSIFKRDGSGSRVFDNRKSGPKRFISNRKSSVKSNGSGNKGMRKSPNKATSR			
DNFKGTGSGSRVFDNRKSGPKRFISNRKSSVKSNGSGNKGMRKSPNKATSR			
TVVYRLNRYKTVSAPVKFNLLNLSLYRKNPTFALNTKRSKPVCTIKSSVPRKIKK			
GLKGSLSKTKRSTSGSKRFLNKFIFLSKINIVTSLNKIPSLPEKQKNTVEVNPESL			
NNATTKKN"			
11241. .14606			
/gene="PF0495w"			
11241. .14606			
/gene="PF0495w"			
/codon_start=1			
/note="predicted by GlimmerM"			
/product="hypothetical protein"			
/protein_id="AAC71888.1"			
/db_xref="GI:3845199"			
/translation="MFALKNTVREGFVNICEFSYLKLLKSNFVTVNLNVTNNEKR			
TNKTYKSKAQSLDFKGLNTHDKLILFNKDPKPKYKCAKEICSAKEVYKLLDEYKKC			
FNYISLCDIIQSVKIFDELDTFTDYFYIEVKNIDKNVNLINKINFYFNKNDITFHR			
ETLKGCKIMSYIHEMNGNELIHFYIFWRNKNDKLLIFFYFNKYFVDFHMYLFN			
HEIYKLLTFNKYLNNSNIPENKNIQEMFENLYVFEIKNEKNYIIMKNKEIYKK			
CPAKPHNVHDIDNEKILNLLYVDNSILDIDINMKLCNLNLLNLIENIETYSKLL			
NFYCYLIRKKGYDNDMTYKUKVEIKATHHILCDTKNLETCSDIDYSTLNSLNNK			
FILNKIIDKNFELYECILLNKIFNFQSLCISLSLKNYYINLRNNVYIVNNY			
LFDIMKFLSYLKNFLGKRITENAVLIHHNDQTNYSNKENIKDIIQKRIEY			
IFYKMEYKDFHKLKDSLDLSIKLSLNTFYKINVSDFYLLFNISCLYLNFLV			
NNSVKYKQTYIYILNDLSFYKYIKNDRTKPKKPFELSSSKELICKNLSVSN			
RYIKHLHEEDFDQDQVCSLTFLNLLRFDKLIHFHYIYNLWCHVYTYNFKCNKL			
INEDIISLLLTCSQFYIENNSDRKELIHLKYNIIDLLKYNLTNTYSISID			
NISKIFISLNSKYTCVENNLLLSQSEFEKVTKTKGGIHMNDLNNDNNSC			
KHYRYIEYKKNFLINLKIIECLINFLYLLKKKTYLYLKQSCLPINLKNIL			
KKLYIANLLYMEYGYCEMLERVLSSHKEONLFYSYNKNVNHKMFDKTLCHSE			
DDYIEMSNTHVLYFDYILKNINSEKQSLNRRNSINDRFIDEIKKKYKLNNTLIK			
NNKLYNEKSNNSNNDLIDKDKNNHNNVEMDLIDNKNENKNIQERQNGNCE			





```

Db 73014 ATATATAAATATATACAAATATATATAGATATATATAAATATATATAAATATATATATAG 73073
Qy 1614 ctcaaatataaacatttaataatgataattctgtaattctcttaattcttcaatttccat 1673
Db 73074 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAG 73133
Qy 1674 atagacacaaacaaataaacatcatcaaatagtttttaacttaacaaacggttaggaa 1733
Db 73134 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAG 73193
Qy 1734 aagttgacctgaactgctgattgacgttggaactgtgcaatgcgaagcgatatttccaa 1793
Db 73194 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 73253
Qy 1794 tatat-----actacatgtagtatttatttatatggaagtcttctcaaaaggtgttgagtg 1848
Db 73254 NNNATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAG 73313
Qy 1849 attgttactgttgaggatgctatttttctcttctgccaataatttttccgagatgg 1908
Db 73314 TAAATATATATATATATATATATATATATATATATATATATATATATATATATATATATAG 73373
Qy 1909 gataactacatactcatgattatgaacgctcaacttttatttgaaaaaacctcctaatacac 1968
Db 73374 TATACATATATATATATATATATATATATATATATATATATATATATATATATATATATATAG 73433
Qy 1969 caaatatgctacatgattcccaaacgtagaccgaattgctctcaatctcaaatcttcaatc 2028
Db 73434 TAAATATATATATATATATATATATATATATATATATATATATATATATATATATATATAG 73493
Qy 2029 aaagtatttaattaccgattgtagaagaaagtttaaccgataataattatcaaaagaagaat 2088
Db 73494 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAG 73553
Qy 2089 aagtcacacagattcttaactcttatttttggtagatgaacatttggtagaataatctcaa 2148
Db 73554 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAG 73613
Qy 2149 aagatgtaactgttttaa 2168
Db 73614 AATATATATATATTTTGAGA 73633

RESULT 11
DMU11584/c
LOCUS      Drosophila melanogaster Oregon-R mitochondrial A-T region.
DEFINITION
ACCESSION      U11584
VERSION      U11584.1 GI:508826
KEYWORDS      mitochondrial DNA; A+T region; tandem repeats.
SOURCE      fruit fly.
ORGANISM      Mitochondrion Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 4601)
AUTHORS      Lewis,D.L., Farr,C.L., Farquhar,A.L. and Kaguni,L.S.
TITLE      Sequence, Organization and Evolution of the A-T Region of
Drosophila melanogaster Mitochondrial DNA
JOURNAL      Mol. Biol. Evol. 11, 523-538 (1994)
MEDLINE      94285822
REFERENCE      2 (bases 1 to 4601)
AUTHORS      Kaguni,L.S.
TITLE      Direct Submission
JOURNAL      Submitted (28-JUN-1994) Laurie S. Kaguni Ph.D. Dept. of
Biochemistry, Michigan State University, East Lansing, MI,
48824-1318, USA
FEATURES      Location/Qualifiers
source      1..4601
             /organism="Drosophila melanogaster"
             /organelle="mitochondrion"
             /strain="Oregon-R"
```

```

/db_xref="taxon:7227"
/dev_stage="embryo"
650..1022
/note="repeat I-A"
/rpt_type=tandem
1023..1360
/note="repeat I-B1"
/rpt_type=tandem
1361..1705
/note="repeat I-C/A"
/rpt_type=tandem
1706..2043
/note="repeat I-B2"
/rpt_type=tandem
2044..2388
/note="repeat I-C"
/rpt_type=tandem
2491..2511
/note="deoxythymidylate stretch"
2512..2648
/partial
/rpt_type=tandem
2649..3112
/note="repeat II-A"
/rpt_type=tandem
3113..3576
/note="repeat II-B1"
/rpt_type=tandem
3577..4040
/note="repeat II-B2"
/rpt_type=tandem
4041..4504
/note="repeat II-C"
/rpt_type=tandem
complement(4565..4585)
/note="deoxythymidylate stretch"

BASE COUNT      2271 a 131 c 74 g 2125 t
ORIGIN
```

Query Match 4.6%; Score 157.6; DB 6; Length 4601;  
Best Local Similarity 44.9%; Pred. No. 9.6e-12;  
Matches 1246; Conservative 0; Mismatches 1469; Indels 57; Gaps 15;

```

Qy 534 aaatttatgtattttctattaaaaagaagaattcttacttacttacttagaattaaa 593
Db 2888 AATTTTATTTAAATTTATTAGATATATAAATATATATAAATATTTTATATATATAA 2829
Qy 594 agctaacgagtgaaacgtagaagcagaatttctaaaaaaaatagtaaacctgcacaaa 653
Db 2828 TATCTATTAATTTATAATTTAGTATATAGTATTTTAAAAAATAATTTATTTTAA 2769
Qy 654 cttatttatgtatataacatatactataaagaacacaaatataatgataaactttaa 713
Db 2768 AAAATTTTATTTTAAATGAAAAATAAATAATATATATTTTCATTATATAAATTTAT 2709
Qy 714 acaaaatttctatgaaattataaaaaaagtcaccttttgacaccttaaaaggttgacaa 773
Db 2708 TAAAAATTTTGTGTTTATTTTAAAAAACATGATTTTATATATAAATTTTATATAA 2649
Qy 774 taaccgtctctccaaaaaaaatacaaacatttataatttctaaaaactatggtgtaatt 833
Db 2648 AATAAATACATTTAAGAAATTTTAAAAAATAATTTATTAATTTTAAAAATAATTTAA 2589
Qy 834 tgcggaatcaaaaaagaagagatttctatcatcaagtttcatttcatgatacaaac 893
Db 2588 TTTCTATATATATATATATATATTAATAATTTCAATAATATATAAATTTATAAATATA 2529
Qy 894 ttcaaatcttcatgtaatttgaaaggaagaaaaataatgaagataaatgtgtttgttctt 953
Db 2528 TAATTAATTAATTTAAAAAATAAATAAATAAATTTTATTTATTTATTTATAA 2469
Qy 954 tat--gttacatttctatggaatatatttcaatacaaaaaaatgtattttaaataatga 1011
```









```
/product="cytochrome c oxidase subunit II"
/db_xref="GI:1166532"
/translation="MSTWANGLDQASPLMEOLIFFHDDHALLILVMTIVLVGYLMFEM
LFNNVNRFLHGLQISIMWITLPAIILLFTALPSLRLLYLLDEINSPVTLKSLGH
QWYSEYSDENNIEFDSYMPITNELMTDGFRLDDVDRNVLPMNSQIRILVTAADVI
HMTVPALGVKVDGTPGRNLQTNFFINRPLGFYGCSEICGHNHSMFPIVIESVPVNY
FKWISSNNS"
tRNA 3768..3838
/product="tRNA-Lys"
tRNA 3840..3906
/product="tRNA-Asp"
CDS 3907..4068
/codon_start=1
/transl_table=5
/product="ATPase 8"
/protein_id="AAC47814.1"
/db_xref="GI:1166533"
/translation="MPQMAPISWLLLIIFISITFILFCSINYSYMPSKSNELKNI
NLSNMNRK"
CDS 4062..4736
/codon_start=1
/transl_table=5
/product="ATPase 6"
/protein_id="AAC47815.1"
/db_xref="GI:1166534"
/translation="MNTNLFVDFDLAIFNLSNWLSTFLGLMIPSIYWLMPRSYNI
MNSLILTHKEFKTLGPGSHNGSTFIFISLFLNPNMGLFPYIFTSTSLTLT
LSLALPLMCLPMYGINHTOHMFALVPOGTPATLMPMVCIEIISNIRPGTLAVR
LTANNIAGHLLTLTGNTGSSSYMLTFLMNAQLLVLESANVMIOYYPVAVLSTL
YSSEYN"
CDS 4736..5524
/codon_start=1
/transl_table=5
/product="cytochrome c oxidase subunit III"
/protein_id="AAC47816.1"
/db_xref="GI:1166535"
/translation="MSTHSNHPFHLDVSPWPLTGAIGAMTTVSGMKVWFHGYDISLF
VLGNTITLTVQWRDVSREGTYOGLHYAVTIGLRWGMILFISELVLFVSPFWAF
FHSLSLPAELGASPPMGIIISFNFOIPLNLTALLASGVTVTWAHHSLSMNNHSQT
TOGLFVLLGYFTILOAYEIEAPFTIADSIYGSTFPMATGPHGIVHLIGTTELLV
CLLRHLNHNHESKNHHFGFEAAWYHFEVDVWVFLYILIIYWWGG"
tRNA 5543..5607
/product="tRNA-Gly"
CDS 5608..5961
/codon_start=1
/transl_table=5
/product="NADH dehydrogenase subunit 3"
/protein_id="AAC47817.1"
/db_xref="GI:1166536"
/translation="MFSIFIALILLITIVMFLASILSKALIDREKSPFCEGFD
PKSSRLPFLRFLLITIFIFDFVEIALILPMIIMKYSNIMIWITISIFILILLI
GLYHEWNOGMLNWSN"
tRNA 5981..6045
/product="tRNA-Ala"
tRNA 6055..6118
/product="tRNA-Arg"
tRNA 6119..6183
/product="tRNA-Asn"
tRNA 6184..6251
/product="tRNA-Ser"
tRNA 6252..6318
/product="tRNA-Glu"
tRNA complement(6337..6401)
/product="tRNA-Phe"
CDS complement(6401..8124)
/notes="TAA stop codon is completed by the addition of 3' A
residue to the mRNA"
/codon_start=1
/transl_except="(pos:complement(6401..6402),aa:TERM)
Query Match 4.6%; Score 157.6; DB 6; Length 19517;
Best Local Similarity 44.9%; Pred. No. 7.7e-12;
```

```
Matches 1246; Conservative 0; Mismatches 1469; Indels 57; Gaps 15;
QY 534 aaattttatgttttttttataaaaaagaagaattcatattactatagaataaa 593
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17804 AATTTTATTTAAATATATAGATATATAATAATATATAATATATATATAA 17745
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 594 agctaagcagttgaaacgtgaaacgagaaatttctctaaaaaaatagaaactgctacaa 553
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17744 TATCTATTAAATATATAATAGTATATAGTATTTTAAAAAATAATTTTAA 17685
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 654 ctattttatgtatataacatatctataaagaacccaataatataataatcattta 713
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17684 AAAATTTTAAAAATGAAATAATAATAATATATTTTCAATATAAAATTTATTTAT 17625
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 714 acaaaattctctgaattataataaaaaagtcacttttgacacttaaaaggtgacaa 773
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17624 TAAAAATTTTCTTTTATTTTAAAAAACATGATTTTATATATAAATTTTATATAA 17565
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 774 taacogtctctccaaaaaaatcaaaacatttataattcttaaaactatggtgtaatt 833
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17564 AAATAATACATTTTAAAGAAATTTTAAAAAATTTATATAATTTTAAATAATTTAAAT 17505
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 834 tgctgaatacaaaagaagaaggatttctatatcataagttcttcatttgcataaac 893
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17504 TTTCTATATATATATATATATATAATATATCAATATATATAAATTTTAAATATATAA 17445
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 894 ttccaatttctgtaatttgaaaggaataaataagataaattgtgtttgtttct 953
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17444 TAATTAATTAATTTTAAAAAATAATAATAATAATAATAATAATAATAATAA 17385
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 954 tat--gttacatttctcgaataatataatcaaaaaaaatgattttaaataatagaa 1011
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17384 TATTAGTATAATAATAATTTTATTTTAAATTTTAAATTTTATTTTAAATTTA 17325
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1012 gagattaccatccaaagtcgaacttatataaaacaagtaataactaaacaacacatg 1071
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17324 AATATTAAATGAATAGATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTAA 17266
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1072 tgatcaaatcaatgacagttttgattctctaaatagaaatgattgagcaaacctcaaaa 1131
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17265 AATTTAAAAATTTTTCATTTTAAATATATATATATATATATATATAATAATTTTATTA 17206
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1132 tgcctcttaggatacaaaacttcccttagcttataaagccgggaggtccaactctc 1191
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17205 TTATTTTAAATTAATTTTATTTTAAATAATTTTATTTTAAAAATAGTTTATTAAGTATAA 17146
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1192 tctcccttgtagacttttggtttccaaatcttcttctccaaaaaatacaataatagttta 1251
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17145 TTT--AATAAATAATTTTAAAAAATAATTTTAAAAAATAATTTTAAAGTTTAAATATAATA 17088
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1252 atgggcataatattggttttaataagtcacatagattttttaggaccatctcctaatcac 1311
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17087 AATTTTATGAATAGGGGAATAAATTTTATTTTTCATTTTACATATATATATATATATATA 17028
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1312 gacaaatatcctaaatgttaacacattttaaacttaaaagatttgcatcacaactcctta 1371
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17027 TATACAATTAATTAATTCAGATTTAGTGATTAATAATAATAATTTTATTTATACACTTATA 16968
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1372 aaatatatatatatatatatatatatatatatatatatataatgaagttatagaaagaa 1431
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16967 ATTTAAATGAAAATTAATATATATATATATATATATATATATATATATATATATATATAT 16908
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1432 cgataaactcttactcaacaattagcccaaaaaaacatccataatgcatttaaaactagga 1491
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16907 AAAAATCAATTTTAAATTTTATTTATTTTAAAAATATTTTAAATTTTAAAAATAAT 16848
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1492 attttaacaaactcaaataggttggtgtagtaaaaaaaacaaatagtagatgacacag 1551
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16847 TTATTTTAAAAATAATTTA---TTATAAAAAATAGTTTATTAAGTATAATAATTTTAAATAATCA 16791
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1552 tacccttaaaatatatactcatatcgaaagtttttaaaatttgcgaaataataacattt 1611
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16790 TTTTTTTAAAAAATAATTTTAAAGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTAGG 16731
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Qy	2671	gggtaggaagagatcgcgtcaacaatcataatgacctccacgcggaagaacttgtcgtttat	2730
Db	15659	TATATAATATATATATATATATATATATATAATAAATAATATTC-----AGATTTA	15615
Qy	2731	acaatttcataacttccggctcggtttatatctccaagtgggtcaaatatattagttt	2790
Db	15614	GTGATTAAATAAATAATTTATTTATATTTTATTTATCGTATAATTTAATAATAACCAAT	15555
Qy	2791	aatacacataacgaggtaataatttggtgactacaattttatcaagtttgggtcaaatagaa	2850
Db	15554	TATTTATATATATATATATATTTTGATTTTTTAAATAAAAATTTTATTTTATTTATTCATA	15495
Qy	2851	acgaacatagtcgtcaaacacgcttcggtgtataccttcatttacgcttaaaaaaacac	2910
Db	15494	AAAAATTTTAATAATTTTTTTTTTAAATAGATTAATAATTTTCATTTTGTAAAAAATAATTTT	15435
Qy	2911	gagaatttgggtcaaatctcaaatatttcataatcatgatataatggaaaattattagatactc	2970
Db	15434	TATAATAAATTTTCATTTTTTTTTTAAAAAATAAATTTTTTTTAAAAATATATATATATA	15375
Qy	2971	tagcgaataatgattgattatgagcttttacaaaaataacgatttttagcattgcaactccct	3030
Db	15374	TTTTTTAAAAAATTTGTAAAAAATAAAGTTTTTAAAAATACCTTAAAAAATAAATTTTTTTTAGT	15315
Qy	3031	tatgtaattcgggtcaaatgttggcatgaagcaagtttgcaccattcaaatttcattta	3090
Db	15314	CATTTAAAGAATAATTAATCATATAATTTTTTAAATTTTTTTTAAAAAATAAATTTTTTTT	15255
Qy	3091	aaatcgtgttgacatacttttaaatactaaatataggaagaagaccaaacattaaattt	3150
Db	15254	AAAAAATAAT--TTTTTTTTAAAAAATAAATAATTTTTTAAAAAAGTGTAAAAATTAATAAT	15198
Qy	3151	agtaagattctaagaacattataagtta-taactataaccacaagaagttgggttta	3209
Db	15197	TTTAATTTTAAAAATTAATAATTTATATATATATTTGATAAATTTATTTTTTTTGGTTA	15138
Qy	3210	gcgtgtgtctctatctgaaacacttgcacaacttaaccatttaattagactcaatgacaat	3269
Db	15137	AAATTTATTTTTTTTTTAAAAAATAAATAATTTTAAAAATTTGTAAAAATTAATAAATA	15078
Qy	3270	taacacacaaat 3281	
Db	15077	AATTTATAAATAAT 15066	
RESULT	13		
PFMAL3P5/c			
LOCUS	PFMAL3P5	86827 bp	DNA INV 04-MAY-2000
DEFINITION	Plasmodium falciparum MAL3P5, complete sequence.		
ACCESSION	AL034556	AL008971	AL008978 AL010141 AL010153 AL010162
VERSION	AL010206	AL010210	AL1391179
KEYWORDS	HTG; centromere; CRP protein; initiation factor E4; Serine/threonine protein phosphatase.		
SOURCE	malaria parasite P. falciparum.		
ORGANISM	Plasmodium falciparum		
REFERENCE	Eukaryota; Alveolata;		
AUTHORS	1 (bases 1 to 86827) Bowman,S., Lawson,D., Basham,D., Brown,D., Chillingworth,T., Churcher,C.M., Craig,A., Davies,R.M., Devlin,K., Felwell,T., Gentles,S., Gwilliam,R., Hamlin,N., Harris,D., Holroyd,S., Hornsby,F., Horrocks,P., Jagsal,K., Jassal,B., Kyes,S., McLean,J., Moule,S., Mungall,K., Murphy,L., Oliver,K., Quail,M.A., Rajandream,M.-A., Rutter,S., Skelton,J., Squares,R., Squares,S., Sulston,J.E., Whitehead,S., Woodward,J.R., Newbold,C. and Barrell,B.G.		
TITLE	The complete nucleotide sequence of chromosome 3 of Plasmodium		
JOURNAL	falciparum		
MEDLINE	Nature 400 (6744), 532-538 (1999)		
REFERENCE	99376085		
AUTHORS	2 (bases 1 to 86827) Bowman,S., Skelton,J., Churcher,C., Lawson,D., Quail,M. and Barrell,B.		

JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 86827)  
AUTHORS Lawson,D., Bowman,S. and Barrell,B.  
TITLE Direct Submission  
JOURNAL Submitted (17-DEC-1998) P.falciparum Genome Sequencing Consortium,  
The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge  
CB10 1SA, UK  
COMMENT On May 5, 2000 this sequence version replaced gi:4493931.  
For more information about this sequence or the Malaria Project,  
see [http://www.sanger.ac.uk/projects/P\\_falciparum](http://www.sanger.ac.uk/projects/P_falciparum).  
FEATURES  
Location/Qualifiers  
1..86827  
/organism="Plasmodium falciparum"  
/strain="3D7"  
/db\_xref="taxon:5833"  
/chromosome="3"  
/clone="MAL3P5"  
gene join(322..603,826..1023,1197..1301,1458..2942)  
/gene="PFC0575w, MAL3P5.1"  
CDS join(322..603,826..1023,1197..1301,1458..2942)  
/note="PFC0575w (MAL3P5.1), Hypothetical protein, len: 689  
aa, possible signal sequence, revised: added new exon 2"  
/codon\_start=1  
/product="Hypothetical protein, PFC0575w"  
/protein\_id="CAB38969.2"  
/db\_xref="GI:7711065"  
/translation="MYLNVIYIISCFILFDCPSFHLKMKYKHNHNMKSVTFLL  
RSPQYKRFRSRKRNKNSFFKKQKPLFLFENLAKGFSGLFGRNOYDOKYIDVOIS  
NNNTRIKQVTHKKSNFEETKENIKQILLHCVSKIDPKIINNLSYLKHPOMSNIT  
VHSILNQISEKVEKKAENYLAHLHLKLDENITLFSWMHIMDFEKKQKVICIRD  
IKSKRRKNLSIYNLICTLIYIETCMCLDIKISHLCIFFFFFFCFGCYNILR  
IYEEVGDLIRKERYNLCEKKIKFHKMDAIKKMEINMKDDDLFYNYHYDELLRC  
FTMKINERNKNIIRSNVDNNDISIDKDMYNNPIDVNTNNTSLDEKIKEOFENP  
DDENLKLDTYEQFLNDNIIKIYEEQDPLYNIDNSNINDNNNTMTMKKHKIK  
DYNLDQGRVKSNDENTSTEHKNTKNTKNGYDFELQONMKNFINKNIDNINSD  
TSNNEINIILKLNQSDSDINLTSDLIYERTRVLWTQKIEYKFKYQDIINQEO  
YPIIRKNTVDLLNGYKIVMSPDVNSLFEKTKIDSIDPNEKKNOMENQKNSKY  
IIMFNVN"  
misc\_feature 604..609  
/note="predicted splice donor sequence for exon 1 of  
PFC0575w (revised)"  
misc\_feature 816..825  
/note="predicted splice acceptor sequence for exon 2  
(revised of PFC0575w)"  
misc\_feature 1024..1029  
/note="predicted splice donor sequence for exon 2 of  
PFC0575w (revised)"  
gene complement(3354..6644)  
/gene="PFC0580c, MAL3P5.2"  
CDS complement(3354..6644)  
/gene="PFC0580c, MAL3P5.2"  
/note="PFC0580c (MAL3P5.2), Hypothetical protein, len:  
1097 aa, possible signal sequence, predicted using  
hexExon"  
/codon\_start=1  
/product="Hypothetical protein, PFC0580c"  
/protein\_id="CAB38968.1"  
/db\_xref="GI:4493932"  
/db\_xref="SPTREMBL:O97257"  
/translation="MLGUKRNVFYLLVSPSLFAYFLKRHKDNENNVTLLNNNDIE  
KIKRIINKSQYIFLLINIDYDIYKNIKLRWLRYFKRRKDKKEYYITNNMVRK  
KRREAYIFIDEQNLNFKFYIYEIVLEYSKYGILSPHLSYLKNIHSEHCNVKYP  
SLUYKLNKLNHNLNKKLVFKQINNEHTQOAPTHTHNNNNNNKKPLDINTHSC  
KNTNYSSTYNNMEKENIYDKYNHNYTEKISYKDNQCQHLTLNMIYLNQTY  
DNICRCLNTNTIYTFMILKAYICKNMEIILYNHIEDMKKINQKNNNTS  
LPKYIYFFFFKKNHYDLFEDQMKNHKKKENDKFNYSNENTHNNIYKNNNTS  
FYDHNSSNRCFSNKLKQDDNTKHIIMGKERYPMKNSDKHKKNNNTGNINIEK  
DQKDLIKLYLNKGLDDIOILNELVMIYMLVLFECSLKISIKKNIHLEKKME  
FDKDNKIYLNADYMMNLRNLKRFKSNERNINFSAPFLLSKNIYFEDEIG  
RSRDNITNNYDKETNTKTTNNNDNDNDNICSNNNDHICSNNNDHICSNN

NNNICSNNNNCSNNNNCSNNNNCSNNKMLDFECQDNKFNNDYTRKKEKKRI  
YELAKIYNNIYDLKGGKXKHQEDNTINLIYIKKKPFWIFYLKNIINKNDTSPIEH  
NNNIVNGDIKNNNIIFKKKYNFLFESSIISYFYKIDYIYNNKLYIYDNLKFCR  
YFKMNEHINRKLKMKRAFHYIYNFQDFIINNTYHIHKKNIHKKIHIHLKQCKDE  
IDVFKDLIYCMINNINNIIFSYIHKVDHNECVYRIFKAYNKILLYEYNLKENIY  
YKNTKTKYLTIVLNNISNDLYPNISYKNIYONKYNKKNFESHIFYSIKNDIHLFL  
LYTORIQNCDFIYSYKKNFNEKPNFLNLYELHYIYVSEKKKKKFFESFTSSP  
YSIDTMVNSFTSYFFSLSLFLFLFHPDMIASYIFFKILYISGLTYYSLYLNNI  
MVCCGPKT"  
gene join(7669..7717,7800..7829,7912..7940,8064..8106,  
8302..8368,8568..8641,8713..8812,8924..9006,9122..9258,  
9369..9505,9613..9838)  
/gene="PFC0581w"  
CDS join(7669..7717,7800..7829,7912..7940,8064..8106,  
8302..8368,8568..8641,8713..8812,8924..9006,9122..9258,  
9369..9505,9613..9838)  
/gene="PFC0581w"  
/note="PFC0581w, hypothetical protein, len: 324 aa,  
revised: new gene prediction, splicing prediction very  
tentative"  
/codon\_start=1  
/product="hypothetical protein, PFC0581w"  
/protein\_id="CAB90285.1"  
/db\_xref="GI:7711066"  
/translation="MKKKKHYESIKFYNDKNKYVLCSDKYIRLINYMIGLHIFRL  
QKIFHYFANDQIASSPFESHKNYRVTKEDIDGIEKCFWNITDYLISESKQDNDF  
NDIKTTVTAMKNKMDOLLTSYNNKIDTVNASFQMASPEYIFINIKFSHRMSPGA  
LKVDDEIVSKNNFSFALSNDNSVTKKYIVDLTLLDNIESKTYNFASVGVVV  
TLKCKELIKWRLKLSKEKYPNMQVMDMKRHHITFTVITNLFLLSLSHRYHDSV  
QNFLKEKNNSDKLQDDIDEDEKYFDEILLREAKKKEEYDKDDEEL"  
misc\_feature 7790..7799  
/note="potential splice acceptor sequence for exon 2 of  
PFC0581w"  
misc\_feature 7830..7835  
/note="potential splice donor sequence for exon 2 of  
PFC0581w"  
misc\_feature 7901..7911  
/note="potential splice acceptor sequence for exon 3 of  
PFC0581w"  
misc\_feature 7941..7947  
/note="potential splice donor sequence for exon 3 of  
PFC0581w"  
misc\_feature 8055..8063  
/note="potential splice acceptor sequence for exon 4 of  
PFC0581w"  
misc\_feature 8107..8112  
/note="potential splice donor sequence for exon 4 of  
PFC0581w"  
misc\_feature 8293..8301  
/note="potential splice acceptor sequence for exon 5 of  
PFC0581w"  
misc\_feature 8369..8374  
/note="potential splice donor sequence for exon 5 of  
PFC0581w"  
misc\_feature 8559..8567  
/note="potential splice acceptor sequence for exon 6 of  
PFC0581w"  
misc\_feature 8642..8647  
/note="potential splice donor sequence for exon 6 of  
PFC0581w"  
misc\_feature 8704..8712  
/note="potential splice acceptor sequence for exon 7 of  
PFC0581w"  
misc\_feature 8813..8818  
/note="potential splice donor sequence for exon 7 of  
PFC0581w"  
misc\_feature 8913..8923  
/note="potential splice acceptor sequence for exon 8 of  
PFC0581w"  
misc\_feature 9007..9012  
/note="potential splice donor sequence for exon 8 of  
PFC0581w"  
misc\_feature 9112..9121



Db 39784 ATTAAATATATTTATTTATTAACGCTTATTTAAATTCATTTAAATTAATATATATATATAT 39725  
QY 2273 taattatgatataaatttgctattcattgcacaaacattactttaaataatcgattt 2332  
Db 39724 TAATTGGAATTTATTTATTTATTAATAAATTTAAATTAATAAACAATTTATTTAT 39665  
QY 2333 ttcattactacaatgttgactagaacaaataacattgattgatgatatacgtcaact 2392  
Db 39664 TAAATTTAAATTTAAATTAATATATATATTAATAATATTTATTTGAAATTAATTTAT 39605  
QY 2393 gaattttcttcgaggagataaattctcaaacatagcaagaatctcataatgatttcg 2452  
Db 39604 TTATTAATTTATTTAAATTAATAATATATATTAATTTATTTATTTAAATTTATTTAT 39545  
QY 2453 tgactacctttagacgaaatttttttaagattcgttaacgtgacttatggtctctgctgt 2512  
Db 39544 ATTATATTAATTAATAATATATATTAATTTATTTATTTAAATTAATTTATTTAT 39485  
QY 2513 ggggggtcaatgcgaataaactcaaatgtatgaggagtcacaaataacacacaaataa 2572  
Db 39484 TTTAATAAATAATATATATATTTATTTATTTTAAATAAATAATATATATTAATTTAT 39425  
QY 2573 aggagcagcaccacaaataaactatagggcagaaatctcttcattggtttaaataaga 2632  
Db 39424 TTTTATTTAAATAAATAATATATTAATTTATTTTAAATAAATAATATATATTA 39365  
QY 2633 ttatccggaagatgaaggactaaaattgaaactgattgggggtag--gaagagatccgctc 2690  
Db 39364 ATTATTTATTTTAAATAAATAATATATTAATTTATTTTAAATAAATAATAT 39305  
QY 2691 acaatcattaaatggctccacgcggaactctgctgttatacaaatcttcacttcggttcg 2750  
Db 39304 ATATTAATTTATTTTAAATAAATAATATATATTAATTTATTTTAAATAAATAAT 39245  
QY 2751 gtcgggtttatctcaaatgggtcaaatgaatttagtttaatacactaacggagaaatt 2810  
Db 39244 AATAATATATTAATTTATTTTAAATAAATAATATATTTATTTATTTAAATTAATTA 39185  
QY 2811 aatgggtactacaattttatcagtttggtgcaattagaaacagacatagtcgtaaaata 2870  
Db 39184 AATAATTAATTTATTTTGTATTTATTTATTTTAAATAATATATATATTTATTTAT 39125  
QY 2871 cgaatcggtgtatacctttattacgttaaaaaataacggaatttctgtcgaattt 2930  
Db 39124 TATATTAATTAATTTATTTATTTTAAATTAATATATATATTAATTAATTAATTTAT 39065  
QY 2931 caaattaattcatgaatatatggaaattattagatactctagcgaataatgattatg 2990  
Db 39064 TTTATTTATTTTAAATAATTTATATTTTATTTTAAATAATA--AATTAATATTTATTA 39007  
QY 2991 agcgttttcaaaaaacagatttttagcattgaaacttcctttatgttaattcggcgaatgt 3050  
Db 39006 ATAATGTTTTTTAAATAATGTAATAATTAATTAATAAATAATGTTATCAATAATTAAT 38947  
QY 3051 tggcagaagaagcaggtttgcaacattaaatttcattcaaaatcgtgtgacatactt 3110  
Db 38946 AGAAAGAAATTAATAAATAATGTAATTAACATATAAATAAATACTATATATATTA 38887  
QY 3111 taaaattcctaataaggagaagaccacaaacattaaatttagaattctaatgaact 3170  
Db 38886 TAAAAATATATATTTATTTTAAACGATGATTATATTTTGTATATTAATAATATTTAT 38827  
QY 3171 ttataagttataactcataa 3190  
Db 38826 ATATTAATAATTAATAAATA 38807

RESULT 14 PFMAL1P3 67970 bp DNA INV 15-DEC-1999  
LOCUS Plasmodium falciparum MAL1P3, complete sequence.  
DEFINITION AL031746  
ACCESSION

VERSION AL031746.9 GI:6594243  
KEYWORDS HTG.  
SOURCE malaria parasite P. falciparum.  
ORGANISM Plasmodium falciparum  
REFERENCE Alveolata: Apicomplexa; Haemosporida; Plasmodium.  
AUTHORS 1 (bases 1 to 67970)  
Bowman,S., Churcher,C., Harris,B., Harris,D., Lawson,D., Quail,M.  
and Barrell,B.  
TITLE Direct Submission  
JOURNAL Submitted (24-SEP-1998) P.falciparum Genome Sequencing Consortium,  
The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge  
CB10 1SA, UK  
COMMENT On Dec 16, 1999 this sequence version replaced gi.5763807.  
For more information about this sequence or the Malaria Project,  
see http://www.sanger.ac.uk/projects/P-falciparum. IMPORTANT: This  
sequence is unfinished and does not necessarily represent the  
correct sequence. Work on the sequence is in progress and the  
release of this data is based on the understanding that the  
sequence may change as work continues. The sequence may be  
contaminated with foreign sequence from E.coli, yeast, vector,  
phage etc.  
FEATURES  
Location/Qualifiers  
1..67970  
/organism="Plasmodium falciparum"  
/strain="3D7"  
/db\_xref="taxon:5833"  
/chromosome="1"  
complement(join(1748..2598,2748..2848,2990..3276))  
/gene="MAL1P3.01"  
complement(join(1748..2598,2748..2848,2990..3276))  
/gene="MAL1P3.01"  
/note="MAL1P3.01, conserved hypothetical protein, len: 412  
aa, similarity: UPF0006 family eg to  
YBL055C/YBL0512/YBL0511, YBF5\_YEAST (418 aa), fasta  
scores: opt: 316, E(): 1.1e-12, (33.2% identity in 271 aa  
overlap)."  
/codon\_start=1  
/product="conserved hypothetical protein, UPF0006 family"  
/protein\_id="CAB63556.1"  
/db\_xref="GI:6594244"  
translation="MKLVFHYIKYINVLFIYSIIFLKSNSLKIYNDLRIYSTVKNKYK  
LQIKRSLKTKNHNIRKMDNESPIDIGSNLTDKMGVYNSKKHENDLQNLNRAK  
NNVDKIIITCTCLAEIDSLKICFTYDEGKFLYLSAGVPTNCYEFIDKNKHEKE  
TIAKEYEEFIKYFKNEQVENSKMENGNKKICDGEKDMNNLEILLENLDTIPGFKY  
NEKDEYLENLKIIKYPNRIVCIGEIGLDFDRFYCSKIQIYFIIPQLKLVQMFN  
LPMFLMRNCSETFEKIVDIYFKERNGVIVHSFTDKEDIVHIIVQYKNLYIGVNG  
CSLSLEINAVAKKIPLLILLLETDAPWCGVKKTHASYEYIKDYTEKRAYTNLKKIKN  
LIKDDNTIFKERNPYNTA"  
complement(2599..2610)  
/note="potential splice acceptor sequence"  
complement(2742..2747)  
/note="potential splice donor sequence, atg/gttaa"  
complement(2849..2861)  
/note="potential splice acceptor sequence"  
complement(2984..2989)  
/note="potential splice donor sequence, aaa/gtaaaa"  
5005..5496  
/gene="MAL1P3.02"  
5005..5496  
/gene="MAL1P3.02"  
/note="MAL1P3.02, hypothetical protein, len: 163 aa,  
contains possible signal sequence"  
/codon\_start=1  
/product="hypothetical protein, MAL1P3.02"  
/protein\_id="CAB63557.1"  
/db\_xref="GI:6594245"  
translation="MKLLNRFVLCPIIIILFFFLNSVVLGNNNRNINFHETENAAK  
AMRKLLSGEINSIKLDNGELKILNDEKHKWDKSDYSFISLEEKESQTDLFR  
KKQINEANTKIIEDRQEFYILNNDENIATRFVLENNFDELYIQSPKQSLIDIQS  
LNN"  
misc\_feature 8020..10389  
/note="possible cent, region of very high [A+T] content"  
14884..20352  
misc\_feature  
gene

```
/gene="MALIP3.03"
14884..20352
/gene="MALIP3.03"
/notes="MALIP3.03, putative ABC transporter, len: 1822 aa"
/codon_start=1
/product="putative ABC transporter"
/protein_id="CAB63558.1"
/db_xref="GI:6594246"
/translation="MTYKENVGISNGKNKKKSCQNTSLNPLSFDWIRLINDLIK
GDIOELNCRNDFVPIYASKLENLRDIEVSEFSEKNSNEHVLHRCNSDASE
KVTNYVYHNILMLFKFRILIIISYLETILVLGGKFIIDYIMRLGOGKIPV
YISLQKDFVSGVLMVIMFFHFFALLHFYFLFTINLKLMLFYLFYKINCSNN
NHLQNPDAFYNTYVTSQSEIDEISDFLSIGKNASSSGSISKNNKNDKNNFVEN
DYILNFKTSYKMEKSDIENRSLPNVNIIMESDVPSTFFVTSCINLFNVFKIF
MSFVYHFKIGSVGIAIMLSIALYSAMILFEPLSKSYLIYRDKRDNNHHVL
KEFKLIKWFESAFKYNIFRKMCKYKIRLYLSNIGVFISSISDIVEVVFIF
YKLDKNNKEIKFTSIMPLIVYIKILISVANFPNLVNMVIGVNIKRLNLTNDH
LYLNDIKNYFRTYRNEVDVYVVDKTLQNETNTHDGTSHNLKHLKNVKNKLTN
MFYFFYFKMYHKNINIKOILSGLLNVDNTNKKICFQEHKSNTSYNNSHIEH
KKEEYENIHNSSTMGNEFKKKNNXYIIKLENCFSGLSYDNKNDCHILKNIFN
LKRNSALIIIGVSGSKAFPHSLTGLDFNTGHLXTENFFKMPILYVQNSWLPWG
NIRSMILFGEYNEPLIYKTYILOSELLNDLSTLPHGDMKYINDDHNLKQKVPICLA
PALYEHVHHKLCIDYEKKLIQNPNEILDKDLNNKNISSYNNKSKLWNYNPFNEN
YLOQLMDNNFYIDLDIPTSISKRIFSNLFCKEDNISFDKDCSFIISMNKS
TLDFLEDLDNVQYVNIFFEQDKTLKYRGNTSEYKNNLTNIESHWGYNLNT
IDYTRIKLFDEVLNHYHSGNMKIYKAYFVKGNTSEFIDSINREYIKMKKKNY
KKEHNNKNDNNNNNNKDDHINTNMNDNHRNYNDINLGNSTODSTPTVSLGNE
YLDQVYSSNDSKEIYKPLYKDPHERFENKSSMPFVKSNNMINDSPKPEYDNSS
FKGSISLETYLWQVGVLLTSTVIFMLISITDEIKFVFLTWMSIISKNKEHSD
TILQKQVILEYFQILPILSVTSIGICFMIYIYNTSAIKVHNHIIISLNAPLYIF
NNLWGLNIINRFIDISAFDGLKRIYKAFPIFRCLISLLIYIMRDCIFFPFV
IILYFVFRFSGCEAQRULSTPLCNITNYSNALSGRNIINIKNTYHLDDVE
HYINFRISYFPFKLIMWASLYIKIFILLTTVIIMPHLYASGIIKLKYEKNYRI
LSTGYSISFASRGLVILKFLCDYTHIEKMCQVORLEPAKTSNKENASMKNEL
NVITQYKKNENISDKSAIYEVKNVSLSSINSQDDSKKYGKIPENYVYSYK
KIPLVNGTYKIDEERSLANNMALKNQIGVSGAGKSTILLSILGLINSOG
KITVEGDIRTYNKGEDSIIIGLAQSFFVYNNIRTFIDPYNFTDDEIVHALKN
GHNKGLNDIYKMHKQDMKSNYKIIQTSRVINQSDNTIILLNDCIRYLSLRYLYN
RIKHYIILIDIEPIFNLNNSVHDLNFLGICAKSFNVIIRNHPNNTVLIISHANT
LSCDDYIVLVRKGBITVRCSDYEDVKTSLSHLEMD"
23896..31533
/gene="rRNA"
/notes="region containing small subunit, 5.8S and large
subunit rRNA genes and spacer regions"
23896..31533
/gene="rRNA"
/complement(join(31966..32476,32675..32775))
/gene="MALIP3.04"
/complement(join(31966..32476,32675..32775))
/gene="MALIP3.04"
/notes="MALIP3.04, conserved hypothetical membrane protein,
len: 203 aa, similarity: P. falciparum chromosome 2,
PF01110W, O96126 predicted integral membrane protein (255
aa), fasta scores: opt: 335, E(): 4.9e-15, (36.1% identity
in 191 aa overlap)"
/codon_start=1
/product="conserved hypothetical membrane protein,
MALIP3.04"
/protein_id="CAB63559.1"
/db_xref="GI:6594247"
/translation="MKASYTFINVTILLFTLLFYNYVDFSTKFFNNNTKIDIN
RFRKRIAESEEQRYPEEDFCILNEELIRPHNDSPYLPEHYENIDRINELINS
TKWKEITKMRQNYEKETDNMNNHNRDFEMHWYKWNIIYLYKVHKLINIITLKDITNPI
HDKBETITTKWQIEDIEYFLENQVLEWRLITLLEFYKNKE"
complement(32477..32486)
/notes="potential splice acceptor sequence"
complement(32669..32674)
/notes="potential splice donor sequence, aag/gtatat"
join(36657..36743,36864..37343)
/gene="MALIP3.05"
join(36657..36743,36864..37343)
/gene="MALIP3.05"
/notes="MALIP3.05, hypothetical protein, len: 188 aa"
```

```
/codon_start=1
/product="hypothetical protein, MALIP3.05"
/protein_id="CAB63560.1"
/db_xref="GI:6594248"
/translation="MRIKMSGIFFIKLLICISFICVFECFKCMISYRKDLLWYSN
CFNYSIDRSLAEGSESSEKTKVDKIPNIELLSLNINYEYKMKETVGFMDNNLN
IANEVKNIHSTNIENIFSLINDSSKSPVLTKLFKEGSIFFHMLNNVPLKLLDLCQ
RNLPHITLGLIVILAAIYVFNFKNEC"
36744..36749
/notes="potential splice donor sequence, aag/gtatga"
36854..36863
/notes="potential splice acceptor sequence"
complement(join(38049..39955,40210..40284))
/gene="garp"
/complement(join(38049..39955,40210..40284))
/gene="garp"
/notes="MALIP3.06, garp, len: 673 aa, similarity: almost
identical to GARP_PLAFF (678 aa), fasta scores: 97.6%
identity in 678 aa overlap"
/codon_start=1
/product="hypothetical garp protein"
/protein_id="CAB63561.1"
/db_xref="GI:6594249"
/translation="MNVFLSYNICILFFVYVTLNFTKFCFSNLLKNQNLNKSFPDS
ITGRLLNTELEKNKDNKNSSETLLKEEKDKDDVPTTSDNKNLNNHNEIISSTDPT
TNIIIVNDKDNENSVDKKDKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK
DKKEENSEVMSLYKTQHKPKNATHEENLYEEMVSEINNAOGLLLSSPYOYRE
OGCGIISVHETSNOTKDNKENSIDKEDHQOEMKLTLDKKERKQEKEMKQOE
KIKKKKKQEKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK
KIDENESTMOQPDSTSETNNEIMVPLSTVDTTPEEHKEGHEKEHEKEHGG
EKHEEHEKEHEKEHEKSKGKKGKKKKKKKKKKKKKKKKKKKKKKKKKKKK
EQUGVETINLEDKEACEQIHTVESRPSLQPKLIDPEQLTMDKSKVVEKNLSIQ
EOLIGTIGRVVVRNDRNHHKMAIEAELQKHVDKEEDKEESKEVEESKEVO
EDDEEEDDEEEDDEEEDDEEEDDEEEDDEEEDDEEEDDEEEDDEEEDDEE
DDAEEDDDDAEDDEDEDEDEDEDEEESKKIRNLRKNNAKI"
complement(39996..40005)
/notes="potential splice acceptor sequence"
complement(40204..40209)
/notes="potential splice donor sequence, aag/gtaaca"
join(45401..46396,46562..50233)
/gene="MALIP3.07"
join(45401..46396,46562..50233)
/notes="MALIP3.07, hypothetical protein, len: 1555 aa"
/codon_start=1
/product="hypothetical protein, MALIP3.07"
/protein_id="CAB63562.1"
/db_xref="GI:6594250"
/translation="MEQIGIANIFTEIGRNVKVRSGPFPIWRKKKFTFEYSKKISNLS
NINIEKCKNCKRINYMDDKKEEFYKFFVEYITSNVKNEFEKAIINTLKNKT
KTLRSRKKYKPKIRIMGELEYIDAINDYKKQFNLYCWSERYIDYKKWLNMKEY
Query Match 4.48: Score 152.2; DB 96; Length 67970;
Best Local Similarity 44.99; Pred. No. 3.2e-11;
Matches 886; Conservative 0; Mismatches 1058; Indels 29; Gaps 7;
Qy 534 aaattatgtatttcttataaaagaagaagaaattctattatagataaa 593
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10493 AAATATATAATTAATTAACATAATATATATAATAAATTTATTAATAA 10434
Qy 594 agctaagcagtgtaaaacgtgaagcagaattcttaaaaaaaatagtaaacgtgacaaa 653
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10433 TAAAGAAAAATAAATAAATAAAGAAAAATTAATAATTTATTAATATAA 10374
Qy 654 ctattatgtgtatataacatactataaagaactcaaatatataataacatttta 713
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10373 ACTTTTAAATTTATTTAAATAAATTAATTAATTAATTAATAAATAAATTTA 10314
Qy 714 acaaaattctatgaataataaaaaaaagtcactttgacacttaaaaggttgacaa 773
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10313 AAATTATATAAAAAAATTAATTAATTAATTAATTAATTAATTTATTTATTA 10254
Qy 774 taaccgtctctccaaaaaaacataatttaattttcctaaacacttggtgaattt 833
```

[illegible]

LOCUS AC005504 10492 bp DNA HTG 01-APR-1999  
DEFINITION Plasmodium falciparum chromosome 12, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 3 unordered pieces.

REFERENCE  
AUTHORS  
1 (bases 1 to 104992)  
Hyman, R.W., Fung, E.L., Qin, F., Tamaki, T., Kurdi, O.B., Conway, A. and Davis, R.W.

**COMMENT**

\* NOTE: This is a 'working draft' sequence. It currently  
\* is not known and their order in this sequence record is  
\* consists of 3 contigs. The true order of the pieces  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will



```

* be preserved. 58642: contig of 58642 bp in length
* 1 58643: gap of unknown length
* 58643 58843: contig of 32169 bp in length
* 58843 91011: gap of unknown length
* 91011 91212: contig of 13781 bp in length.
* 91212 104992: Location/Qualifiers
FEATURES
source
1..104992
/organism="Plasmodium falciparum"
/db_xref="taxon:5833"
/chromosome="12"

```

BASE COUNT	44286 a	9326 c	9564 g	41411 t	405 others
ORIGIN					

Query Match 4.4%; Score 151.4; DB 60; Length 104992;  
Best Local Similarity 46.0%; Pred. No. 3.7e-11;  
Matches 904; Conservative 0; Mismatches 1021; Indels 40; Gaps 10;

[illegible][illegible]



Qy	2406	agggatataattctcaaacatagcaagaatctcataataatgttt	2450
Db	72798	ATTTTATTATTATAAAATTATAATATTTTTTATTTTAAATATTT	72754

Search completed: July 13, 2001, 00:36:41  
Job time: 9809 sec

**THIS PAGE BLANK (USPTO)**